Sevenoaks School

Prized writing

International Baccalaureate Extended Essays 2015

Rory Alexander - Art

Arvind Arora - Geography

Ale Baranowski - Biology

Olivia Brandon - English

Loris Gliner - Mathematics

Emma Harrington - World Studies

Claudia Hockey - History

Max Kitson - Economics

Laura Lau - Chemistry

Abie Witts - Spanish

- Prized Writing -

Sevenoaks School
International Baccalaureate
Extended Essays 2015

Introduction from John Sprague, Director of IB, Sevenoaks School

Every year in February each of our Lower Sixth students embarks on a journey which will be the one of the biggest challenges of their young academic careers: the IB Extended Essay. It is undoubtedly a daunting task: a 4000-word independently researched essay on a topic of their own choice, often in a subject that they've only been studying for six months. Every year, however, all of our students reach their goal having produced a remarkable piece of original research. I say 'independent' and genuinely mean it. The IB is pretty clear in its expectations: supervisors are to spend no more than five hours over a six-month process working directly with the student and at no point are we allowed to edit the student's writing or compel them into any particular direction. The students choose and develop their ideas; we help them bring them to fruition. It is without a doubt the most difficult academic work in which they will have engaged to date. It is also the element of the IB which seems to most prepare them for the rigour and independence of undergraduate study.

Again and again we hear from our visiting alumni that it was the Extended Essay that prepared them most for university research and gave them a clear advantage among their undergraduate peers. The collection of essays contained in this, our first edition of Prized Writing, are genuinely that – prized. We, their supervisors and teachers, think they stand out solely on the basis of the originality of their topics, the levels of commitment and enthusiasm of the authors and the demonstrated ability of these students to overcome whatever difficulties they encountered in the research and writing process. The essays were purposely chosen before they received any external marks from IB examiners. We are, undoubtedly, proud of each of the 212 essays submitted to the IB in May 2015 and would stand by any one of them, but these seemed to be first among equals, so we share them with you.

They are included unedited, in exactly the way that they are submitted to the IB, so there may be some idiosyncrasies within, but this is what makes them so special. They've been written by real students doing their absolute best, and we applaud them.

Contents

Do the design and appearance of the Jin Mao Tower draw more heavily on Chinese or Western architectural influences?

Rory Alexander - Visual Art

Following a late subject change, I realised a new Extended Essay in the Visual Arts was an opportunity for me to explore an area that I have a genuine passion for: architecture. Incorporating my interest in Chinese culture, I decided to examine the architectural influences on the design of the iconic Jin Mao Tower in Shanghai. After attending a summer course in China, I was able to visit Shanghai's historical Longhua Temple complex as well as the Jin Mao Tower itself, to understand the historical Chinese architectural motifs evident in this modern International Style skyscraper. My research ultimately afforded me a fascinating insight into contemporary issues of national identity in a modern and globalised China. On reflection, using my Extended Essay to explore my own personal interests allowed me to thoroughly enjoy an experience I had initially approached with dread.

Supervisor: Charley Openshaw

Rory's essay is of course appropriately weighty in its rigorous analysis of the aesthetics and construction of the buildings he examines. It is however, not this aspect that strikes the reader, rather the coolly expressed passion for his subject that runs through all of his writing. He is very evidently awed by the ambition, scale and design of the Jin Mao Tower in Shanghai and the intense nature of his site visits is eloquently expressed. It is this rich balance of a personal exploration into the potential of creativity and a disciplined analysis of monumental architectural form that make this such a persuasive essay.

Contents

Abstract

This essay will answer the question of whether the design and appearance of the Jin Mao Tower in Shanghai draw more heavily on Chinese or Western architectural influences. The introduction examines some of the key characteristics of both Western and Chinese architectural culture, establishing a basis upon which to evaluate the influences that acted upon the design and appearance of the Jin Mao Tower.

Using evidence gathered first hand from site visits to the Jin Mao Tower itself as well as the Longhua Temple complex in Shanghai, the investigation is divided into three sections: an analysis of the materials and techniques used in the structure, the composition of the building, and its surface and interior architectural decoration. Investigating these aspects allows the essay to develop a sense of the interwoven influences evident in the building's design and the resultant emotional response when seeing the building in person. An overall understanding of the Jin Mao Tower as relying on Western architectural principles concerning skyscrapers is diversified by examining evidence of the tower's decidedly Chinese identity, which conveys an atmosphere of peace and within the urban environment.

Further to this, the essay examines the extent to which Chinese motifs in the tower can be regarded as genuine and not superficial, raising questions about the tower's intended identity. Thus, the essay arrives at a conclusion that the architects behind the tower acknowledge the paradox of representing historical Chinese architectural traditions in a modern and predominantly Western form of building. The building can therefore be seen as being a contemporary creation in itself, but creating a dialogue with ancient architectural traditions in China. As a part of the Lujiazui district master plan, the Jin Mao Tower thereby comes to represent China's past, with two successively larger skyscrapers portraying China' present and future.

Figure 1: Jin Mao Tower, Shanghai

Introduction

Completed in 1999, the Jin Mao Tower in Shanghai is widely described as recalling 'the ancient form of the Chinese pagoda.'[1] This popular view of the tower as a direct architectural reference to historical Chinese culture is particularly interesting given its simultaneous status as the 'crown jewel'[2] of Shanghai's modern, forward-looking economic centre of Lujiazui.

Twenty-three years after the death of Chairman Mao Zedong, the tower was completed at a time when economic liberalisations had begun to take effect on China, as demonstrated by an average yearly economic growth rate in the 1990s of over 10%[3]. Thus, as the first super-tall skyscraper to be built in Shanghai's Lujiazui district (now one of the most recognisable skylines in the world), the Jin Mao Tower can be seen as a building that had to represent the changing identity of a modern China.

In the search for internationally renowned architects to design the tower, the Chicago-based firm SOM was eventually awarded the project, led by architect Adrian Smith. It is somewhat intriguing that a firm involved so heavily in the design of American skyscrapers was given the task of creating a uniquely Chinese one. This irony provides the basis of an investigation into how the Jin Mao Tower was conceived as a modern skyscraper while embodying a Chinese identity.

In order to evaluate the extent to which the Jin Mao Tower draws inspiration from Chinese and/or Western architectural influences, it is crucial to outline some of the key characteristics of both types of architecture. Though attempts to define an entire architectural culture may be regarded as somewhat futile, by delineating specific characteristics of architecture in both cultures, it is possible to understand which culture has been most influential in the design and appearance of the Jin Mao Tower.

In this essay, the term 'Western' will be used to describe both European and North American architecture, due to their strongly interconnected nature. Several architectural developments extending from the late Nineteenth Century must be considered when describing the traits of Western architecture. Chicago is widely credited as the birthplace of the skyscraper thanks to the innovation of the loadbearing steel frame, first seen in Burnham and Root's Reliance Building, 1890-4. This development allowed buildings to reach higher, and soon a new theme emerged within urban architecture- verticality. Sullivan's Guaranty Building in Buffalo, 1894-5 (Fig. 2), was the first great example of emphasis on verticality, with its strong vertical lines and recessed horizontals fitting into a

[1] Riley, T., Nordenson, G., 2003, Tall Buildings, New York, The Museum of Modern
[2] Dupre, J., 2008, Skyscrapers- A History of the World's Most Extraordinary Buildings, New York, Black Dog & Leventhal Publishers, pg121
[3] Guiness, P., 2011, Geography for the IB Diploma- Patterns and Change, Cambridge, Cambridge University Press, pg129

structure divided clearly into three parts: base, shaft and top. From this prototype developed the evolving form of the skyscraper, with a focus on Art Deco emerging in the 1920s and 1930s and encapsulated by the soaring, streamlined shape of New York's Chrysler Building, 1930-1 (Fig. 3).

Figure 2: Guaranty Building, Louis Sullivan, 1894-5

Figure 3: Chrysler Building, William van Alen, 1930-1

In the early to mid Twentieth Century, architects became obsessed with the theme of Utopia, merging political ideologies with designs for entire cities. Architects involved in this trend became known as the Futurists, and one good example can be seen in Antonio Sant' Elia's Terminal for Aeroplanes and Trains with Funicular, 1914 (Fig. 4). One movement that was not confined to the drawing board was Modernism. Pioneered by Mies van der Rohe, this movement held a belief in itself as a 'democratic answer to social crisis'[4], conveyed through clean-cut grids, largely minimalist in nature, personified famously in Mies van der Rohe's Seagram Building, New York (Fig. 5) Within this fast-paced evolution, the architectural firm SOM, from the 1950s onwards, became heavily involved in the development of the glass-box skyscraper ideology, and the resultant dominance of International Style.

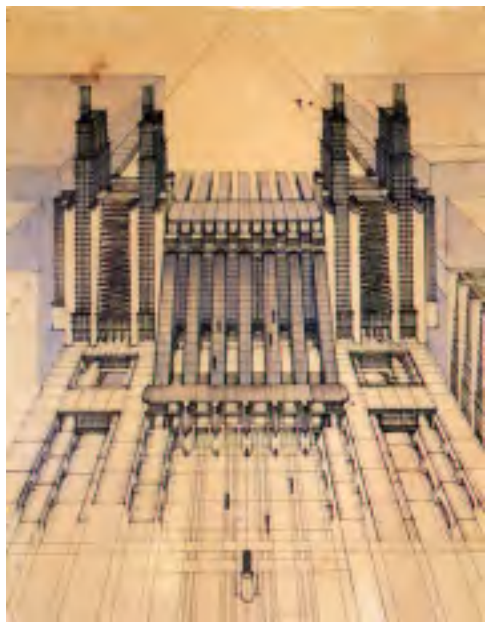


Figure 4: Terminal for Aeroplanes and Trains with Funicular, Antonio Sant' Elia, 1914

[4] Hughes, R., 1991, The Shock of the New, London, Thames & Hudson, pg167

Figure 5: Seagram Building, Ludwig Mies van der Rohe, 1958

Figure 6: Longhua Temple Pagoda

Meanwhile, Chinese architecture can be seen as having a more constant focus on the technical genius of structures alongside beautifully crafted ornamentation. Chinese architecture more consistently spans the different building uses that a society requires, from residential dwellings, to Temples. Perhaps the most widely recognisable symbol of Chinese architecture is the ancient Chinese pagoda. The pagoda (wooden or stone), often regarded as one of the earliest forms of skyscrapers, is distinctive in the way it comprises of a large column-like shape, narrowing at each progressive protruding set of wooden or stone eaves. The form of the pagoda gives insight into other traits of Chinese architecture, the exposed wooden eaves of Chinese roofs, for example, or the significance of colour and handcrafted decoration.[5] The Longhua Temple Pagoda (Fig. 6), Shanghai is a strong example of a wooden pagoda. Less tangible characteristics include the use of auspicious numbers in design, and the significant meanings attached to certain shapes and other motifs. Thus, Chinese architectural history provides an overwhelming depth to be appreciated in a modern world, and hence recalled in contemporary Chinese architecture.

This essay will examine three key aspects of the Jin Mao Tower's design and appearance, using evidence gathered from both secondary sources and from a

[5] Qijun, W., 2011, Discovering China: Chinese Architecture, Shanghai, Shanghai Press and Publishing Development Company, pg9

site visit to the tower itself. Firstly, the materials and techniques used in construction must be considered. The composition of the building itself must then be investigated, to understand the precise means by which the building emanates the form of the traditional Chinese pagoda. Finally, the surface and interior architectural decoration used will give further insight into the interaction of Chinese and Western influences in the finishes of the final product. This essay will thereby answer a question that, during an on going Chinese construction boom, is significant in showing how a modern China can recall its own heritage through architecture.

Materials and Techniques

In terms of structure, as a skyscraper, it is largely inevitable that the Jin Mao Tower draws more heavily on American architectural influences. Like any modern skyscraper, the tower was constructed using concrete, steel and glass. An octagonal central core of reinforced concrete walls acts as a central load-baring structure. Eight super-frame columns are then spaced around the perimeter of the floor plan, linked to the core by outrigger trusses at 3 points on the twenty-sixth, fifty-third and eighty-seventh floors of the tower.[6] This allows unimpeded floor spaces enclosed by a glass 'curtain-wall'. The John Hancock Centre, Chicago, 1969 (Fig. 7), also designed by SOM has a very similar floor plate in terms of the arrangement of the core and the perimeter load baring columns. The Jin Mao's design can therefore be seen to mirror that of the majority of Western skyscrapers, and uses all of the same materials.

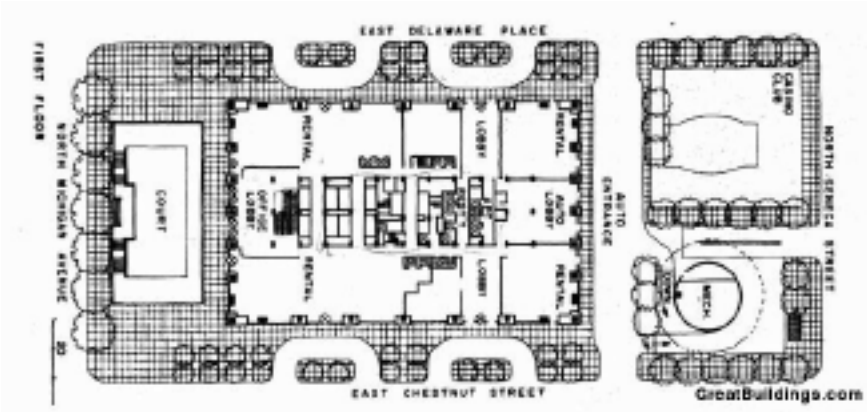


Figure 7: John Hancock Centre ground floor plan, SOM, 1969

Though Western engineering provides the key structural principles for the modern skyscraper, the ancient Chinese pagoda, as an early form of skyscraper, can also offer innovative solutions for weight distribution within a skyscraper. While the design of the Jin Mao Tower's load-baring structure is inherently based on Western architectural principles, the system of outrigger trusses seems to parallel the cantilever rafters present in ancient wooden pagodas. Pagodas such as the Longhua Temple Pagoda are designed to suspend a central pillar-shaped mass by the 'upward force of all the cantilever rafters'[7]. Though this principle itself is not used in the tower, the outrigger trusses are somewhat reminiscent of it, and the arrangement of the structure as an embedded concrete and steel tube with the vast empty space of the thirty-eight floor Grand Hyatt Hotel lobby in the uppermost section of the tower inverses the Chinese architectural concept of a suspended column, with the rooted column of the tower instead housing a cavernous empty space (Fig. 8).

[6] Riley, T., Nordenson, G., 2003, Tall Buildings, New York, The Museum of Modern Art, pg138
[7] ibid.

One of the most evident materials used in the tower is glass. Each level in the Jin Mao Tower is clad in glass, largely drawing on its functionality in letting in natural light and allowing for panoramic views.[8] These uses make it unsurprising that the Shanghai Foreign Trade Centre Company desired a glass-clad skyscraper, as any other business would. Though this may be seen as a Westernising factor in the building's design, the innovative decoration of this glass façade, as discussed later, creates a means of dialogue with Chinese traditional architecture.

The materials and techniques used in the construction of the building's structure therefore draw most heavily on Western architectural influences, while the reversal of the structural principles of the ancient Chinese pagoda makes reference to the architectural traditions that are most evident in historical Chinese architecture. A sense of the modern structure acknowledging its own context within a country with such a rich history is thereby achieved.

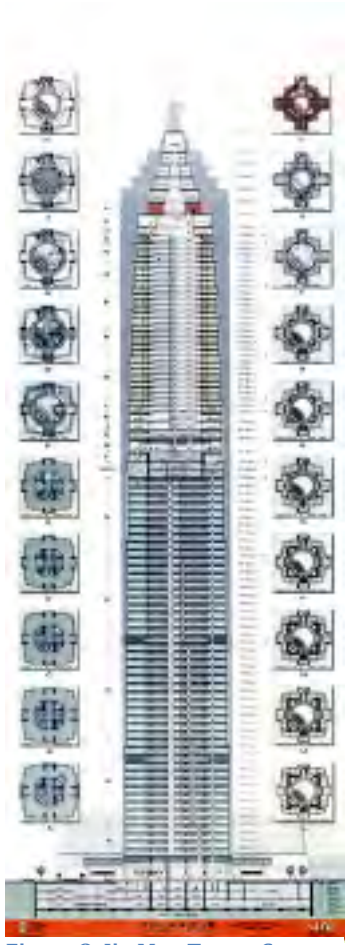


Figure 8: Jin Mao Tower Cross Section showing open lobby section in the top part of the building

[8] http://www.bbc.co.uk/news/magazine-27501938

Composition

In this essay, the topic of composition refers to the physical form of the Jin Mao Tower, while drawing on less tangible aspects of architectural influence, including the significance of numbers in the tower's design, contributing to an overall sense of the tower's architectural theme. The physical form of the building relates strongly to its effect on the viewer, so investigating composition can be seen as crucial in understanding the response of the viewer, and the identity that the tower conveys.

One of the key traits of the modern skyscraper is its progression from base, to the rising shaft and eventually the top or crown of the building. Though nowadays, due to the onset of Modernism and the more recent popularity of International Style (the glass-box skyscraper), the base and crown of a skyscraper have perhaps become less strongly emphasised, the Jin Mao Tower reverts to this early Western form of skyscraper composition, but in a way which is decidedly Chinese in inspiration.

Figure 9: Base of the Jin Mao Tower

The most widely noted aspect of the Jin Mao Tower's composition is its pagoda-like form. The influence of the ancient Chinese pagoda on the Jin Mao Tower contributes to all 3 parts of the building's form: base, shaft and crown. The tower sits on a stone-clad base (Fig. 9) that can be linked to the base of numerous historical Chinese constructions. The base of the tower inclines slightly, becoming more slender as it rises. This type of base is a common sight in historical Chinese city gate designs, which throughout history 'chiefly served the purpose of defence.'[9] The 12th Century artist Zhang Zeduan captures the strong shape of city gates in his painting, 'Along the River During the Qingming Festival' (Fig. 10). This stone silhouette conveys a sense of strength and stability, ideal for the historical defensive outer gates of a Chinese city. The base of the Jin Mao Tower itself acts as a gateway into the building, and its silhouette invokes a sense of the stability and majesty of the skyscraper above.

[9] Qijun, W., 2011, Discovering China: Chinese Architecture, Shanghai, Shanghai Press and Publishing Development Company, pg19

Figure 10: Section of 'Along the River During Qingming Festival', scroll painting, Zhang Zeduan, 12th Century (Song Dynasty)

The main shaft of the skyscraper embodies the form of the ancient Chinese pagoda. The series of thirteen rhythmic setbacks (Fig. 11), which progress further up the tower, give the building its recognisable pagoda silhouette. The outward sloping of the façade beneath each setback recalls the eaves visible beneath the setbacks present in both stone and wooden pagodas. The building even has the traditionally desirable odd number of setbacks[10], with thirteen such sets of eaves, relating directly to the stone Tie Ta Pagoda (Fig. 12).

Figure 11: Jin Mao Tower with setbacks

Figure 12: Tie Ta Pagoda, Kaifeng, China

[10] http://architecturemps.com/wp-content/uploads/2013/09/mc_conference_day_kirsten1.pdf

The progressive increments of the tower also serve to exaggerate its height as seen from street-level. From beneath, the decreasing distance between each setback is less easy to perceive, as seen during the site visit. Instead, the sense of perspective which one has when looking up at the building acts to suggest that it is even taller than it is, as the increasing number of setbacks near the top, can seem to be of an equal distance apart, instead of having a decreasing distance in between one another. This clever progression highlights one of the key architectural themes of the Jin Mao Tower: verticality. Here, the line blurs between the Chinese and Western architectural influences, as architect Adrian Smith uses a Chinese motif to fulfil what would traditionally be classed as a Western theme of verticality. This sense of height is extended by the use of a single continuous vertical stainless steel-clad protrusion that runs up the centre of each of the 4 faces of the tower (Fig. 13). As the floor plate decreases in area further up the building, this sharp metallic line cuts straight upward, from street-level to crown. Thus, emphasis of the vertical shows a typically Western desire to build tall emerging through the use of Chinese motifs, resulting in a strong soaring effect when seen in person.

Figure 13: Emphasis on verticality

The crown itself is a culmination of the tower's pagoda outline. The floor plate withdraws from the perimeter in a series of heavily ornamented steps that culminate in the sharp point of the spire. The point of the spire itself does not bare any close similarities with the tops of ancient Chinese pagodas. Chinese pagodas can be topped with very complex spires, often more sculpted and curved than angular and jagged, as demonstrated by the ornamentation of the crown of the Longhua Temple Pagoda in Shanghai (Fig. 14). The progression of strong geometry present in the Jin Mao's crown may be more closely linked to

the Art Deco movement. One travel guide even goes so far as to describe the tower as 'New York's Chrysler Building (Fig. 3) reimagined as a 21st-Century Pagoda.'[11] Though the jagged, metallic quality of the crown alongside its illumination at night (Fig. 15) certainly recalls the crown of the Chrysler Building, the tower cannot be described as solely Art Deco. Instead, Art Deco elements contribute to the tower in creating a dialogue with Art Deco buildings on the Bund (Fig. 16), on the opposite side of the Huangpu River.

Figure 14: Longhua Pagoda Crown Ornamentation

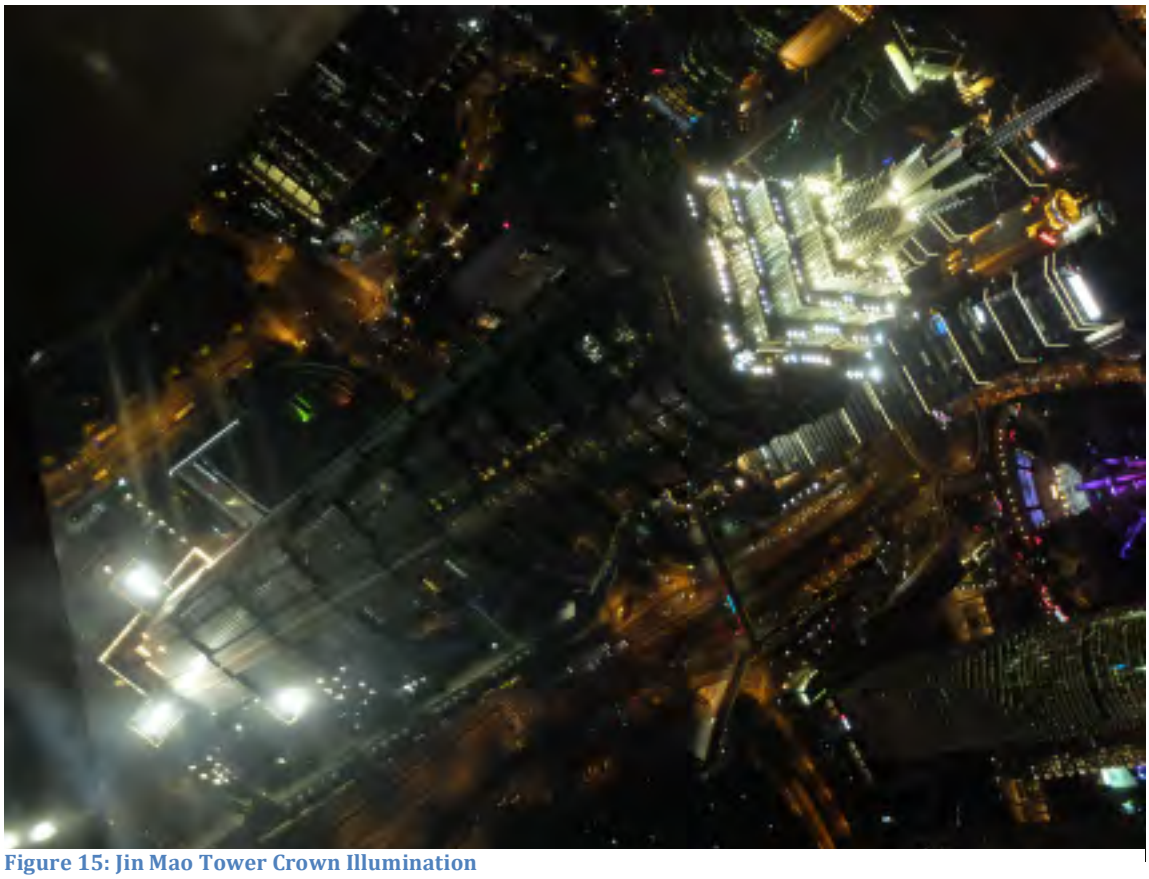


Figure 15: Jin Mao Tower Crown Illumination

[11] http://travel.nationalgeographic.com/travel/city-guides/shanghai-walking-tour-1/

Figure 16: Art Deco Buildings overlooking the Huangpu River, The Bund, Shanghai

The composition of the Jin Mao Tower was particularly heavily influenced by one requirement of the clients, China Shanghai Foreign Trade Centre Company. The requirement was that the building's design should incorporate the number eight, an auspicious number in Chinese culture, having a similar pronunciation to the word for 'wealth'.[12] The building is located at 88 Century Boulevard in Lujiazui, but the references to that number are far more deeply woven into the architecture of the tower itself. The Jin Mao Tower also has eighty-eight floors, and its octagonal concrete core eight sides, both auspicious and another link to the many eight sided ancient Chinese pagodas. Moreover, the distance between the successive set-backs moving up the tower decreases by one eighth with each new set-back. Hence, good fortune is seen to be brought to those who inhabit the tower, hence the translation of Jin Mao to mean "Golden Prosperity".[13] The building does not, however, abide completely by traditional rules of Feng Shui, perhaps as superstition has, since the rise of the (atheist) Communist Party, been officially frowned upon in China, where during the Cultural Revolution, any such 'mystic belief' was 'banned'.[14]

[12] http://www.travelchinaguide.com/intro/lucky-number8.htm
[13] http://architecturemps.com/wp-content/uploads/2013/09/mc_conference_day_kirsten1.pdf
[14] http://www.nytimes.com/2013/05/11/world/asia/feng-shui-grows-in-china-as-officials-seek-success.html?_r=0

The smaller podium building of the Jin Mao Tower (Fig. 17) lies directly adjacent to the main tower, on its Western side. The inclined stone walls of the tower's base encircle the podium building as well, conveying a greater concept of strength and rigidity for the entire building (tower and podium). This rigid exterior is accented by a curving roof that once again references the upturned eaves of a traditional Chinese building. The curving metallic planes of this roof structure are less directly representative of Chinese architecture, and more of a type of Post-Modernist ornamentation for the podium. The sweeping curves can be seen to mirror the ornamental crown of New York's Post-Modernist Sony Tower (Fig. 18). In this way, in person, the podium appears as a kind of caricature, showing the building to acknowledge the paradoxes of modern and historical China. The podium thus stands to justify the tower itself, acknowledging that its reference to Chinese architecture should not be considered truly authentic.

17: Podium Building (left)

Figure 18: Sony Building, New York

The somewhat bizarre metal ventilation pipes (Fig. 19) that jut out of the Eastern wall of the podium seem to consolidate this idea, contrasting with the ornate surrounding surface decoration. The tubes' functionality may acknowledge a basic need for utility within the building or in a wider sense suggest the lack of a perfect modern Chinese image. Either way, this jarring glimpse of the building's services echoes the far more boisterous Lloyd's Building (Fig. 20), in London. While Rogers' Lloyd's Building embraces functionality in the most extreme sense, the subtle hint to it in the Jin Mao Tower allows a more scaled back insight into the dynamics of a monument which stands in a modern China and refers to a historical China.

Figure 19: Ventilation pipes

Figure 20: Lloyd's of London, Richard Rogers

Surface and Interior Architectural Decoration

The ornamentation of the façade of the Jin Mao Tower can give a significant insight into the influences that acted upon the architects responsible for its design. The Jin Mao Tower's façade consists of steel, aluminium and glass. Unusually, for a Western-designed skyscraper, the ornamentation of the façade of the tower protrudes 46cm from the surface of the perimeter glass wall.[15] The vertical and horizontal steel tubes that protrude from the façade create a strong grid pattern that repeats itself in the silvery reflections of the glass windows (Fig. 21). The use of a repetitive grid pattern on the tower's façade can be seen to recall the 'modular grid' of Western skyscrapers, seen as the 'face of modernity'[16]. The façade of Mies van der Rohe's Seagram Building, New York (Fig. 5) is the embodiment of the modular grid of modernism. However, upon comparison of the two facades, it is clear that the Jin Mao Tower's only loosely refers to the modular grid form, developing its own intricacy and thus its own identity, as it intensifies over different portions of the façade. Thus, the façade is difficult to class as a part of a defined Western architectural style.

Figure 21: Jin Mao Tower Facade

The reflective windows of the tower, coupled with the steel that encircles the building, give the Jin Mao Tower a very distinct colour. The silvery-bronze colour of the façade adds to the sense of the tower retaining its own identity, despite having design traits that associate the building with any number of Western

[15] Dupre, J., 2008, Skyscrapers- A History of the World's Most Extraordinary Buildings, New York, Black Dog & Leventhal Publishers, pg121

skyscrapers. Parts of the façade, through the webbing of steel, convey a metallic quality that can be paralleled with the metallic quality of Western buildings incorporating Art Deco style into their design. The intricate printed patterning appearing on the surface of the windows themselves further contributes to this sense of a link with Art Deco. However, one should hesitate in describing the tower as Art Deco, thanks to its manipulation of its exterior metalwork. The intensification of metalwork at the top of each increment of the tower instead of recalling the metallic Art Deco skyscraper, recalls, through its darker colour, ad intricacy, the wooden structure visible in the upturned eaves of an ancient Chinese pagoda. Thus, the surface decoration persists in creating a Chinese style of ornamentation, despite open reference to specific periods of Western architectural style.

Figure 22: Circular doorway at the Longhua Temple Complex

Chinese thought and historical architectural practices also directly contribute to the building's exterior and interior. The circle is a shape often repeated in historical Chinese architectural designs. Within Chinese culture, the circle is traditionally representative of 'unity'[17] and a 'round (circular) heaven'.[18] The Temple of Heaven in Beijing is a Temple complex devoted to the worship for the Heavens, with the 'circular mound altar'[19] building being the focal point of the entire complex. Here, the circle is used to symbolise the Heavens. Interestingly, the circle is also a shape used throughout the Longhua Temple Complex in Shanghai. Longhua, a Buddhist Temple complex is likely to associate the shape with the concept of unity. In the complex, circles are used for doors (Fig.22) and gateways, suggesting a harmony that is achieved by entering into the refuge of

[17] http://www.illuminantpartners.com/2012/04/23/more-than-just-a-circle-and-squareshapes-chinese-culture/
[18] Qijun, W., 2011, Discovering China: Chinese Architecture, Shanghai, Shanghai Press and Publishing Development Company, pg91
[19] ibid.

the Temple complex. In the Jin Mao Tower, the circle is also used in the doors that lead into the tower from three sides (Fig. 23). In the Jin Mao Tower, the use of circles may draw on the idea of the Heavens in the sense that the tower reaches upwards for the Heavens, alternatively, it could relate to securing a sense of harmony and peace within the building. Overall, circles create a persona that removes the building from the air of functionality associated with most office towers. The motif of the circle is also used nearer the top of the Jin Mao Tower, in the shape of the thirty-eight-floor hotel lobby (Fig. 24). Further to this motif, the spiralling balconies of the lobby can be directly associated with the spiral's meaning in Chinese thought, as 'suggest(ing) the universe and eternity'[20], having a mesmerising effect on the viewer. Thus, these aspects of Chinese architectural though are strongly integrated into the design, creating an overall sense of calm solitude, away from the busy city environment. It could be argued, however, that the use of these motifs is purely superficial.

Figure 23: Circular entrance

Figure 24: Hotel Lobby with spiral effect

[20] Pridmore, J., 2008, Shanghai: The Architecture of China's Great Urban Centre, New York, Harry N. Abrams, Inc., pg68

Conclusion

This evidence must now be drawn back to the question of whether the Jin Mao Tower relies more heavily on Chinese or Western architectural influences. Overall, the design of the Jin Mao Tower can be seen to have a basis in Western architectural principles and styles, which is manipulated by the architect in order to create an aesthetically unique building that relates itself directly to its Chinese architectural context. As a skyscraper, it is largely inevitable that the tower relies heavily on Western architectural principles. However, through parallels with the form of the ancient Chinese pagoda, and through the use of a nuanced collection of Chinese motifs, the Jin Mao Tower retains a uniquely Chinese identity, establishing an atmosphere closely linked to that of a Temple complex within the city environment.

Though some may criticise the building's design for referring too much to Western architectural principles and styles, the identity of the tower can be seen as appropriate in a modern China. The façade's patterning recalls the modernist grid that later developed into International Style. Meanwhile, many point out the similarities that the tower shares with its Western Art Deco counterparts from the 1920s and 1930s, but as Charles Jencks argues, it is dangerous to 'dismiss Jin Mao as Art Deco'.[21] Rather, we should appreciate the dialogue that the tower creates with Shanghai's own Art Deco history, while sensing other styles at play in the building. The similarities to be drawn with Post-Modernism also create an interesting point underlined by the metal vent which juts out of the podium; perhaps the architects are accepting that an aim to recreate Chinese architecture perfectly is flawed. It might be better to think of the building as acknowledging its contextual past whilst adapting to represent the requirements of a modern China: functionality and success.

Adding weight to this concept is the current plan for the on-going development of Shanghai's Lujiazui district. Including the Jin Mao Tower, three stunning super-tall skyscrapers will dominate this skyline, with Jin Mao standing next to two taller buildings: Shanghai World Financial Centre (SWFC) and Shanghai Tower (Fig. 25). In this group, the Jin Mao Tower is said to represent China's past, SWFC China's present, and Shanghai Tower, China's future.[22] Thus, Shanghai's skyline begins to show its own unique Utopian ideals of China's identity in the modern world.

[21] Jencks, C., 2012, The Story of Postmodernism- Five Decades of the Ironic, Iconic and Critical in Architecture, Hoboken, John Wiley & Sons, pg120
[22] http://events.cornell.edu/event/gensler_shanghai_tower

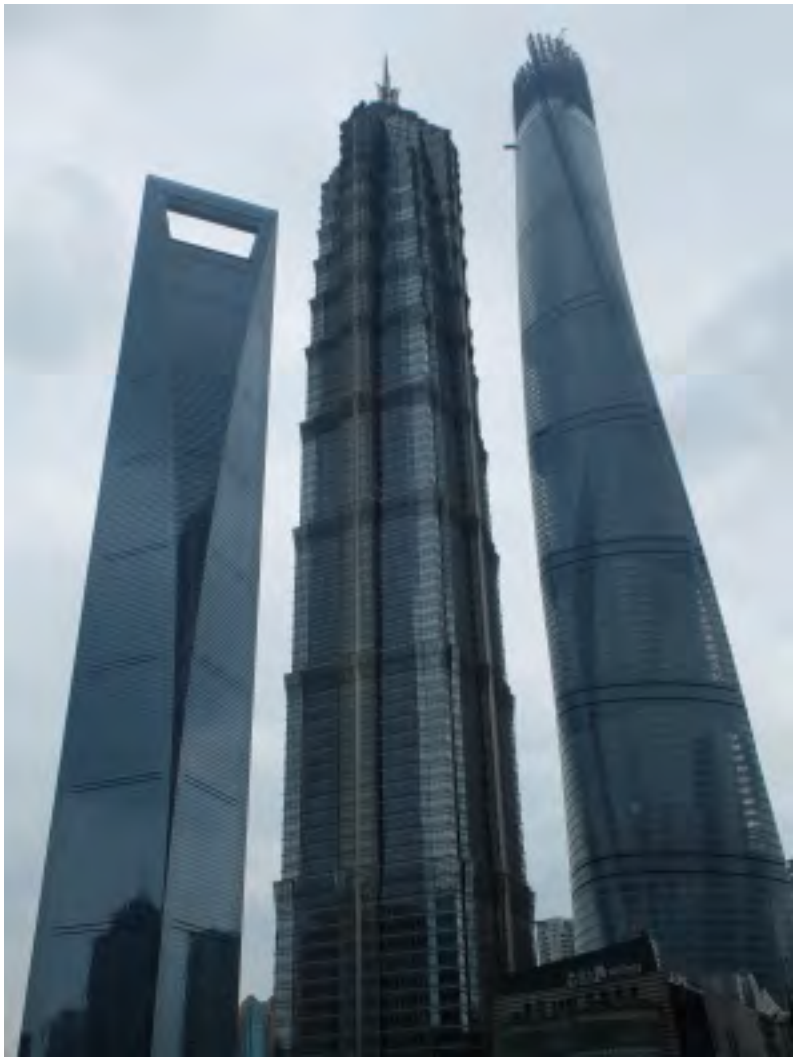


Figure 25: (left to right) SWFC, Jin Mao Tower, Shanghai Tower

Bibliography

Books:

Riley, T., Nordenson, G., 2003, Tall Buildings, New York, The Museum of Modern Art

Dupre, J., 2008, Skyscrapers- A History of the World's Most Extraordinary Buildings, New York, Black Dog & Leventhal Publishers

Guiness, P., 2011, Geography for the IB Diploma- Patterns and Change, Cambridge, Cambridge University Press

Hughes, R., 1991, The Shock of the New, London, Thames & Hudson

Qijun, W., 2011, Discovering China: Chinese Architecture, Shanghai, Shanghai Press and Publishing Development Company

Pridmore, J., 2008, Shanghai: The Architecture of China's Great Urban Centre, New York, Harry N. Abrams, Inc.

Jencks, C., 2012, The Story of Postmodernism- Five Decades of the Ironic, Iconic and Critical in Architecture, Hoboken, John Wiley & Sons

Websites:

http://www.illuminantpartners.com/2012/04/23/more-than-just-a-circle-and-squareshapes-chinese-culture/, accessed 26/08/14

http://architecturemps.com/wp-content/uploads/2013/09/mc_conference_day_kirsten1.pdf, accessed 26/08/14

http://www.bbc.co.uk/news/magazine-27501938, accessed 26/08/14

http://www.nytimes.com/2013/05/11/world/asia/feng-shui-grows-in-china-as-officials-seek-success.html?_r=0, accessed 26/08/14

http://events.cornell.edu/event/gensler_shanghai_tower, accessed 26/08/14

Image Sources:

Figure 1: Author, taken 16/08/14

Figure 2: http://www.historic-structures.com/architecture/chicago_school/index.php, accessed 25/08/14

Figure 3: http://www.greatbuildings.com/cgi-bin/gbi.cgi/Chrysler_Building.html/cid_1165529717_Chry_Closeup.html, accessed 25/08/14

Figure 4: http://www.artribune.com/2013/04/un-secolo-di-citta-nuove/2-513/, accessed 25/08/14

Figure 5: http://archinect.com/news/tag/64597/design-observer/15, accessed 25/08/14

Figure 6: Author, taken 15/08/14

Figure 7: http://www.greatbuildings.com/cgi-bin/gbc-drawing.cgi/John_Hancock_Center.html/John_Hancock_Plaza.jpg, accessed 25/08/14

Figure 8: http://skyscraper.org/EXHIBITIONS/CHINA_PROPHECY/jin_mao.php, accessed 25/08/14

Figure 9: Author, 16/08/14

Figure 10: http://en.wikipedia.org/wiki/Architecture_of_the_Song_dynasty#mediaviewer/File:Bianjing_city_gate.JPG, accessed 25/08/14

Figure 11: Author, taken 16/08/14

Figure 12: http://www.visitourchina.com/guide/kaifeng.htm, accessed 25/08/14

Figure 13: Author, taken 16/08/14

Figure 14: Author, taken 15/08/14

Figure 15: Author, taken 14/08/14

Figure 16: Author, taken 14/08/14

Figure 17: Author, taken 16/08/14

Figure 18: http://structurae.net/structures/sony-tower, accessed 26/08/14

Figure 19: Author, taken 16/08/14

Figure 20: http://modernarchitecturelondon.com/pages/lloyds.php, accessed 26/08/14

Figure 21: Author, taken 15/08/14

Figure 22: Author, taken 16/08/14

Figure 23: Author, taken 16/08/14

Figure 24: Author, taken 16/08/14

Figure 25: Author, taken 16/08/14

To what extent does The Netherlands follow the Kuznets Curve of Environmental Degradation?

Arvind Arora - Geography

The hardest part of the EE was actually choosing among all the interesting topics on my mind; it took me a good two before I narrowed down on investigating the changes in wealth and environmental degradation in The Netherlands. Gathering, organizing and analyzing all the data took up much of the time of the process, however, as each data point was analyzed, and as each graph or choropleth were plotted, it was amazing to see the resulting trends - expected or not expected. Writing the Extended Essay was time consuming, however, having chosen a topic which really interested me, I found the task as a whole really enjoyable and it taught me many skills which I have used throughout the IB and I'm sure which I will use into the future as well.

Supervisor: Oliver Russell

The analysis and discussion of this essay goes far beyond the expectations for an Extended Essay. Arvind took on a difficult challenge in taking on a broad concept - 'environmental sustainability' and managed to dissect and evaluate this in an appropriate manner. To do this he investigate a transect from West to East across The Netherlands, assessing different economic and environmental factors to analyse how environmental sustainability is affected by affluence. His discussion of market forces and justification of his results is superb, helping to clearly explain a large data set in a concise manner. This work goes beyond what one would expect from many undergraduate students. Arvind's attitude, diligence and independence in producing this work, for me, makes this essay stand out as a most impressive piece of work..

Acknowledgements

I would like to thank all my friends and family for their support over the period when this extended essay was written; I especially would like to thank my sister Aditi Arora for all her moral support and tips throughout the summer!

More over, I would like to thank my brilliant extended essay supervisor Mr. Olly Russell. All his guidance and geographical wisdom was invaluable as I searched for an appropriate and interesting topic to investigate; I wouldn't have been able to do this without him!

Contents

Abstract

This essay seeks to answer the question: To what extent does the Netherlands follow the Kuznets Curve of Environmental Degradation?

The data was collected regarding 6 different environmental factors - the average number of cars per person, the average household waste per kilogram per person, the number of motorbikes per 1000 people, the average natural gas usage per household in meters cubed, the average electricity usage per household in kilowatt per hour and lastly, the average delivery of electricity per person.

A transect was drawn, going from west to east Netherlands traversing 26 different cities. The data was collected from the CBS StatLine census information belonging to the Dutch Government. Using this information provided by census, the environmental effect (degradation) of each of the 6 different environmental factors was determined and the changes in the variable was correlated to distance from the sea; the distance from the sea was highly linked to wealth as cities further east had a lower income compared to cities on the west.

The table below shows what values would be required for a city to be the **least** environmentally degrading in respect to the 26 cities along my transect (see methodology) -

Average number of cars per person	Low – 0.08
Average household waste per KG per person	Low – 368 KG pp./ year
Number of motorbikes per 1000 people	Low – 34
Average natural gas usage per household in meters cubed	Low – 1250 m^3 ph./ year
Average electricity usage per household KwH	Low – 2800 KwH pp./ year
Average Delivery of Electricity KwH	Low – 1323 KwH ph./year

Using the data, a series of graphs were created, as well as choropleth maps and for each variable, the spearmans rank correlation coefficient was worked out as well. Conclusions were then made about the changes in the environmental degradation seen regarding affluence. A Kuznet Curve based on each environmental factor was then drawn and compared with all the other environmental variables.

It can be seen that in general, the most and least affluent cities in The Netherlands had the highest environmental degradation, whereas the middle income regions had the lowest environmental degradation. The Netherlands was found not to follow the **published** Kuznets Curve at all, however, The Netherlands did follow a reversal or mirror image of the published Kuznets Curve.

List of Figures

List of Tables

1. Introduction

The term ‘environmental degradation’, and subsequent questions of sustainability, have been recognised as key phrases shaping international and national discussions. Particularly due to the context of climate change and its environmental effects, issues regarding environmental degradation and sustainability are at the forefront of politicians, activists and increasingly, local populations. Environmental degradation is defined as the ‘deterioration in environmental quality from ambient concentrations of pollutants and other activities and processes such as improper land use’ (OECD 2014). Sustainability is more difficult to define, however, the Brundtland commission’s widely cited definition is adapted to provide the best example; ‘Sustainability creates and maintains the conditions under which humans and nature can exist in productive harmony, in order to meet the needs of current generations without compromising the ability of future generations to meet their own needs’ (Adapted from EPA 2014, Brundtland 1987).

There has already been research surrounding these topics, in particular work which provides comparisons between countries. Moreover, there has been a focus on developing nations (Shandra, London & Williamson 2003); as a result, this essay moves away from the bulk of the research done and tackles the issues of degradation and sustainability by conducting an indepth study on The Netherlands – with comparisons at a regional scale.

The Netherlands is a low-lying country in Europe (Figure 2) with a population of approximately 16.8 million (CIA World Factbook 2014). Moreover, it has a population area slightly less than twice the size of New Jersey (NationMaster), with an average of 405.5 people/km^2; and as a result, The Netherlands is the 63rd most populated country and the 24th most densely populated country in the world (CIA World Factbook). Therefore, pressure on resources is incredibly high; for example, the country has the highest oil prices in the European Union (Figure 1).

Figure 1 Map of Europe showing the variations in oil prices per country

Figure 2 Map of Europe and the world showing the location of The Netherlands

In order to analyse how sustainable and environmentally friendly The Netherlands is, the Kuznets curve of Environmental Degradation has been used. An approach such as this regarding The Netherlands is novel, and as a result may provide an interesting understanding of Dutch environmental impacts. The Kuznets curve is ‘a hypothesized relationship between various indicators of environmental degradation and income per capita’ (Stern 2003). The curve shows the expected level of environmental degradation level by correlating it with changes in income per capita (Figure 3).

Figure 3 The Kuznets Curve of Environmental Degradation

2 Disparities in The Netherlands

2.1 Economic Disparities

The Netherlands has a Gini coefficient of 30.9 (CIA World Factbook, 2007), However, Van Bavel (Bijlo 2014) described the country as very economically uneven. He went further and stated that the distribution is 'more uneven than in most European countries, even more uneven than in The United Kingdom and just as unevenly distributed as in the United States' (Adapted from Van Bavel in Bijlo 2014). It has been concluded that the three richest Dutchmen hold more wealth than approximately half of all Dutch households put together; whilst, the bottom sixty percent of Dutch households hold less than one percent. (Bijlo 2014).

Figure 4 Graph showing the changes in the Gini coefficient for The Netherlands from 1977 till 2011

2.2 The Kuznets Curve

The environmental sustainability of an area can be dependent on changes in affluence; therefore, with the reported spread of wealth in The Netherlands as mentioned above, it can be predicted that the environmental sustainability will change throughout the country. To preliminarily evaluate the application of the Kuznets curve to The Netherlands, and therefore the feasibility of this essay; cities with the lowest, the median, and highest income per-capita (Table 1) out of 26 cities lying on a defined East to West transect (see methodology) will be taken and then correlated to a few environmental factors.

2.2.1 Preliminary evaluation of the Kuznets Curve

.

Figure 5 Map showing the location of the 3 cities on the defined East to West transect

Name of City	Averaged Income (€)
Wassenaar	51,900
Utrecht	31,400
Haaksbergen	27,600

Table 1 Showing the average income for three cities in The Netherlands

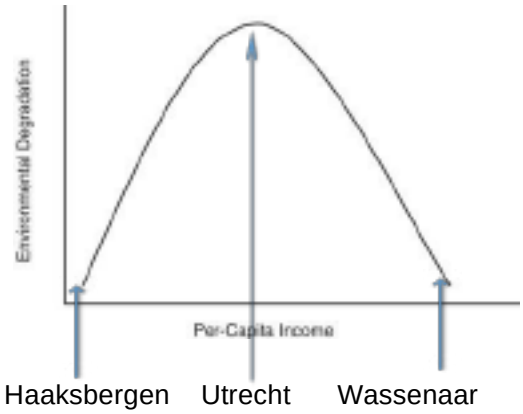


Figure 6 The Kuznets Curve showing the relative positions of the three chosen cities

If the Kuznets curve is applicable to The Netherlands, the level of environmental degradation for each of these three cities should follow the trend of the curve.

Name of City	Household Waste per person (kg)	Natural Gas Usage per person (m^3)	Electricity Usage per person	Total	Maximum	Env. Degradation from maximum possible. value
Wassenaar	610	3600	2650	6860	7317	94%
Utrecht	524	1250	**2800**	4574	7317	63%
Haaksbergen	**667**	**3850**	2300	6817	7317	93%

Table 2 Showing the values for the chosen environmental factors for the three cities.

Using household waste, natural gas usage and electricity usage, the three cities were equated to the highest possible environmental degradation values for each of the three categories (shown in bold on Table 2). Wassenaar is 95% of the maximum values for all three categories, with Haaksbergen being 94% and Utrecht 63%. These values result in a curve as shown in Figure 7.

Figure 7 shows an inverse trend in environmental degradation versus per capita income as compared with the original KC. It shows that in The Netherlands, the most and least affluent cities along the E-W transect have the highest environmental unsustainability, whereas the middle city has the lowest. As a result of this interesting finding, this work will continue to

study the Kuznets Curve and the effect of this inverse trend further by using a wider variety of economic data for all 26 cities along the transect.

Figure 7 The trend of environmental degradation and per-capita income found from the three cities

3 Aims

The aim of this essay will be to determine the extent to which the Kuznets Curve of Environmental Degradation is applicable to The Netherlands. If it is applicable, then using certain statistical analyses will provide statistical proof. However, if it is not, then the reason for the difference will be questioned and alternate factors will be examined in order to explain why.

3.1 Hypothesis

“The Kuznets Curve for The Netherlands will be the inverse of the published curve, whereby the most and least affluent areas will have the highest environmental degradation, and the middle income cities will have the lowest ”.

4 Methodology

To carry out the investigation, it would have been very tedious to look at all 408 cities in The Netherlands, therefore a transect going from the East to the West of The Netherlands was delineated; the cities used were those that lay on this transect.

The transect, shown in black in Figure 8, resulted in 26 cities included in the investigation and evaluated:
Alphen aan den Rijn, Amersfoort, Apeldoorn, Barneveld, Berkelland, Bodegraven, Bronckhorst, Brummen, De Bilt, Ede, Haaksbergen, Leiden, Leiderdorp, Leusden, Lochem, Nieuwkoop, Rijnwoude, Soest, Utrecht, Voorschoten, Voorst, Wassenaar, Woerden, Zeist, Zoeterwoude, Zutphen.

Figure 8 Map showing the 26 cities on the West to East transect

4.1 Data Collection

Secondary data from the Centraal Bureau voor de Statistiek (Central Bureau of Statistics) was collaborated for all 26 cities. Moreover, for each, data was collected on 6 different environmental factors, and 2 economic factors in order to evaluate degradation and affluence.

The environmental factors:
1. The average number of cars per person
2. The average household waste per kilogram per person
3. The number of motorbikes per 1000 people
4. The average natural gas usage per household in meters cubes
5. The average electricity usage per household in kilowatt an hour
6. The average delivery of electricity per person per 1000 meters cubed

The economic factors:
1. The average household income
2. The number of immigrants per 1000 inhabitants

To obtain all the data required the 'CBS StatLine' website was used, which is the Dutch data bank for statistics. All 26 cities were added to those required by using their 'gemeente code' which is their municipality code. The CBS StatLine database then relayed the data categories available for all 26 cities, from which the six, varied environmental factors and the two economic variables were chosen.

For each of the variables, the data was downloaded and imported into Excel. The values for each factor were then ranked from large to small - ranking the largest as number 26 and the smallest as rank 1, this was repeated for the 6 factors. Furthermore, each of the city's distance from the sea was calculated, going from 0 km to 186 km, and ranked from closest to furthest. **The distance from the sea is used as a proxy for affluence because the further one goes from the sea, the less affluent each city becomes**. Once this was completed, a Spearman's rank correlation coefficient calculation was obtained for the factors and they were subsequently correlated with the distance from the sea. Once this calculation was complete and the values were plotted on the degrees of freedom graph, and it was able to evaluate the statistical significance the correlation between the two factors. As well as this, the variables were plotted on a bar graph, adding a trend line in order to better visualize the results and make predictions. Moreover, all the data was added into a Google Documents Fusion Table application, in conjunction with all the polygon geometry data for the 26 cities; allowing for a Choropleth map to be created for the variables and hence showing the changes in the factor going from East to West along the transect.

4.2 'The Most Sustainable City'

In addition to the analyses conducted mentioned above, the environmental factor data for the 'perfect city' was calculated.

From the six factors chosen, the lowest value from each one represented the most environmentally friendly area. Then the values for the 'ideal' and most sustainable city for each factor were taken and added together. Subsequently, each city's performance was evaluated as a percentage with regards to how close they came to the perfect city.
Table 3 shows the method used to create the perfect city. First the lowest values from each category were highlighted and the sum of these values was taken. Lastly, the actual values for each city were added together and then this total was divided by the lowest possible total of the perfect city. As a result, how well various cities' performed, in relation to the perfect city as the gold standard, was evaluated in terms of a percentage value (table 4).

Table 3 Showing the values and calculation used to create the perfect city.

The percentage shows how far off each city from being the most sustainable. Wassenaar was the furthest from the most sustainable city, whilst Utrecht was the most sustainable from the 26 cities chosen.

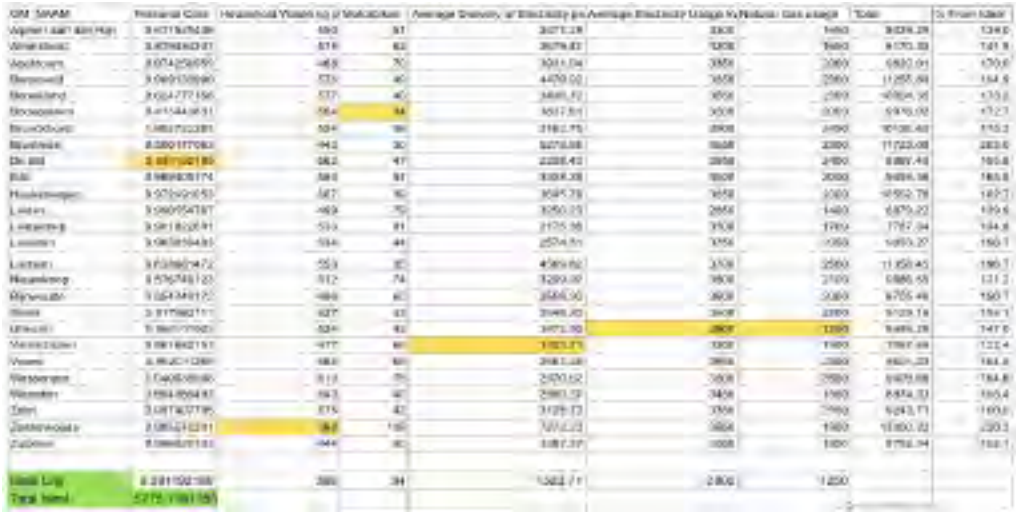


Table 4 Showing the percentage distance from each city to the perfect city

Finally, the values for most sustainable environmental city and most sustainable economic city were added together (31404.25 + 7575.796) and the summed values for each city's environmental and economic data were analyzed as a percentage of the above total.

The final percentage value obtained shows by how many percent above or below the city is from an ideal ('most sustainable') city. After which, the Spearman's rank correlation coefficient and degrees of freedom test for this value was calculated.

5 Results

5.1 Economic Data Results

5.1.1 Average Household Income

GM_NAAM	Averaged Income	Rank	Distance from the sea	Rank Sea		Difference in Ranks	Ranks D^2
Alphen aan den Rijn	32400	17	25	7		-10	100
Amersfoort	33200	18	79	15		-3	9
Apeldoorn	29700	9	121	18		-9	81
Barneveld	28800	6	101	17		11	121
Berkelland	28600	1	171	25		24	576
Bodegraven	30600	24	32	8		-16	256
Bronckhorst	31000	10	150	23		13	169
Brummen	27700	3	131	20		17	289
De Bilt	29800	6	61	12		6	36
Ede	29600	8	121	18		10	100
Haaksbergen	27600	2	186	26		24	576
Leiden	31400	13	11	3		-10	100
Leidendorp	34700	21	15	5		-16	256
Leusden	34800	22	81	16		-6	36
Lochem	31100	10	156	24		14	196
Nieuwkoop	31000	10	35	9		-1	1
Rijnwoude	31600	16	20	6		-10	100
Soest	34200	20	71	14		-6	36
Utrecht	31400	13	52	11		-2	4
Woerdenstein	38700	25	9	2		-23	529
Veenst	28300	4	131	20		16	256
Wassenaar	51900	26	0	1		-25	625
Woerden	33600	19	40	10		-9	81
Zeist	36600	23	69	13		-10	100
Zoeterwoude	31400	13	14	4		-9	81
Zutphen	28300	4	140	22		18	324
					Sum D^2		5036
							-0.722353616239

Table 5 Showing the data for household income

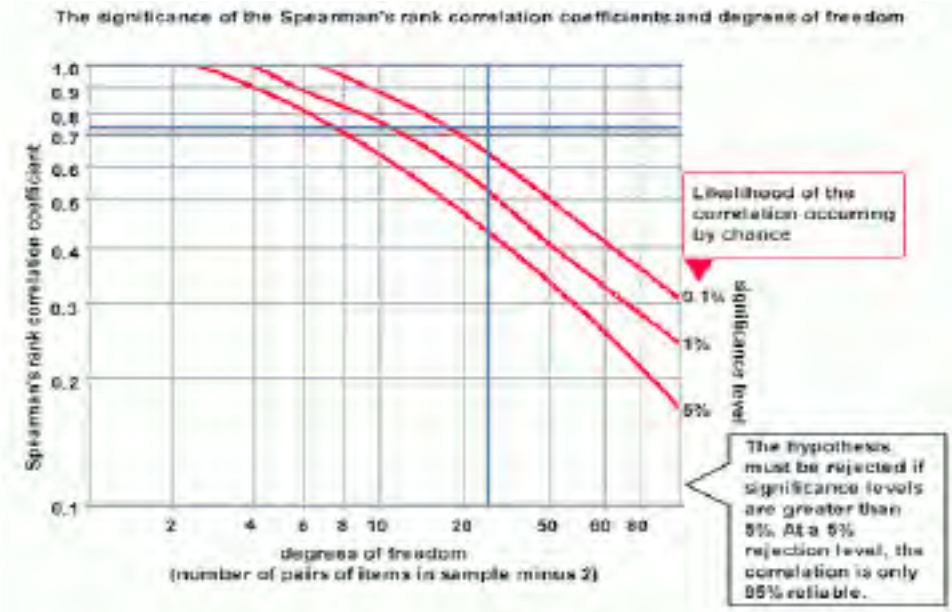


Figure 9 The degrees of freedom graph for household income

For correlating household income and distance from the sea (West to East), the Spearman's rank value of **R=-0.722** was calculated which indicates that there is a very strong negative correlation. The degrees of freedom graph illustrated that the correlation is more than 99.9% reliable and is therefore significant. This shows that with distance from the sea, the average income decreases.

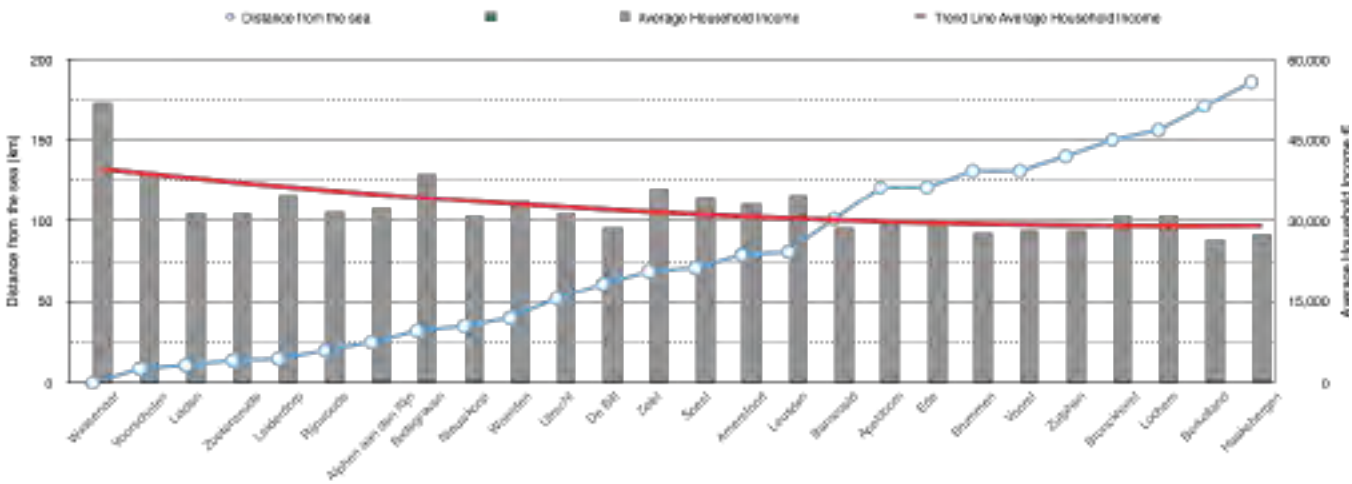


Figure 10 a bar chart showing the trend between household income and distance from the sea (affluence)

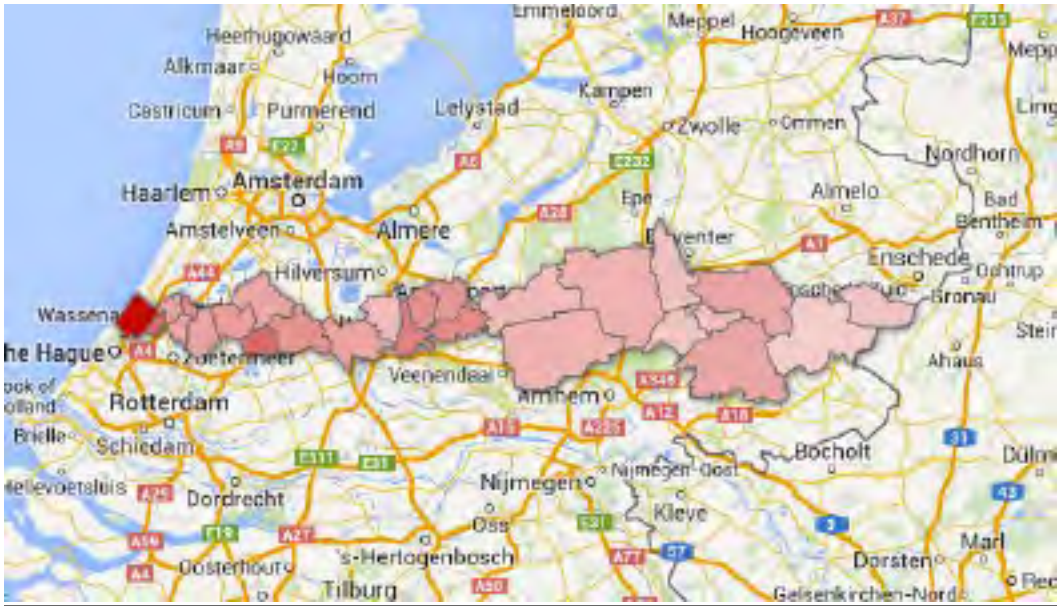


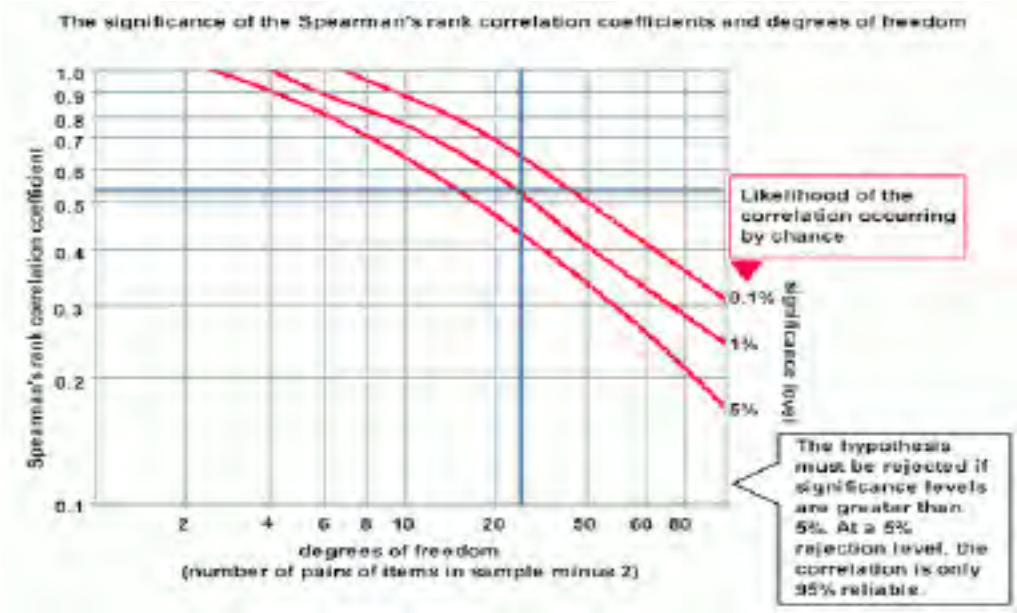
Figure 11 a Choropleth map showing changes in household income along the transect

5.1.2 Number of immigrants per 1000 inhabitants

GM_NAAM	Migration	Rank	Distance from the sea	Rank Sea		Migration D	Migration D^2
Alphen aan den Rijn	5.5	19	25	7		-12	144
Amersfoort	8	22	79	15		-7	49
Apeldoorn	4.3	14	121	18		4	16
Barneveld	3.5	11	101	17		6	36
Berkelland	2.2	5	171	25		20	400
Bodegraven	6.7	20	32	8		-12	144
Brandehorst	3.8	12	150	23		11	121
Brummen	2.1	4	131	20		16	256
De Bilt	2.7	8	61	12		4	16
Ede	4.2	13	121	18		5	25
Haaksbergen	1.8	1	188	26		25	625
Leiden	17.4	25	11	3		-22	484
Leiderdorp	4.9	17	15	5		-12	144
Leusden	2.8	9	81	16		7	49
Lochem	2.4	6	156	24		18	324
Nieuwkoop	3.4	10	35	9		-1	1
Rijnwoude	2.1	3	20	6		3	9
Soest	4.9	18	71	14		-4	16
Utrecht	14.4	24	52	11		-13	169
Veenschoten	7.9	21	9	2		-19	361
Voorst	1.9	2	131	20		18	324
Wassenaar	24.6	26	0	1		-25	625
Woerden	4.4	15	40	10		-5	25
Zeist	8.8	23	69	13		-10	100
Zoeterwoude	2.6	7	14	4		-3	9
Zutphen	4.6	16	140	22		6	36
						Sum D^2	4508
							-0.541196581196581

Table 6 (Above) Showing the data for the number of immigrants per 1000 inhabitants

Figure 12 (Below) showing the degrees of freedom graph for the no. of immigrants per 1000 inhabitants



For correlating immigration and distance from the sea (West to East), the Spearmans rank produced a value of **R=-0.541** which indicates that there is a strong negative correlation; and on the degrees of freedom graph we can see that the correlation is more than 99% reliable,

and therefore the correlation is significant. As one goes away from the sea, the number of immigrants per 1000 inhabitants decreases.

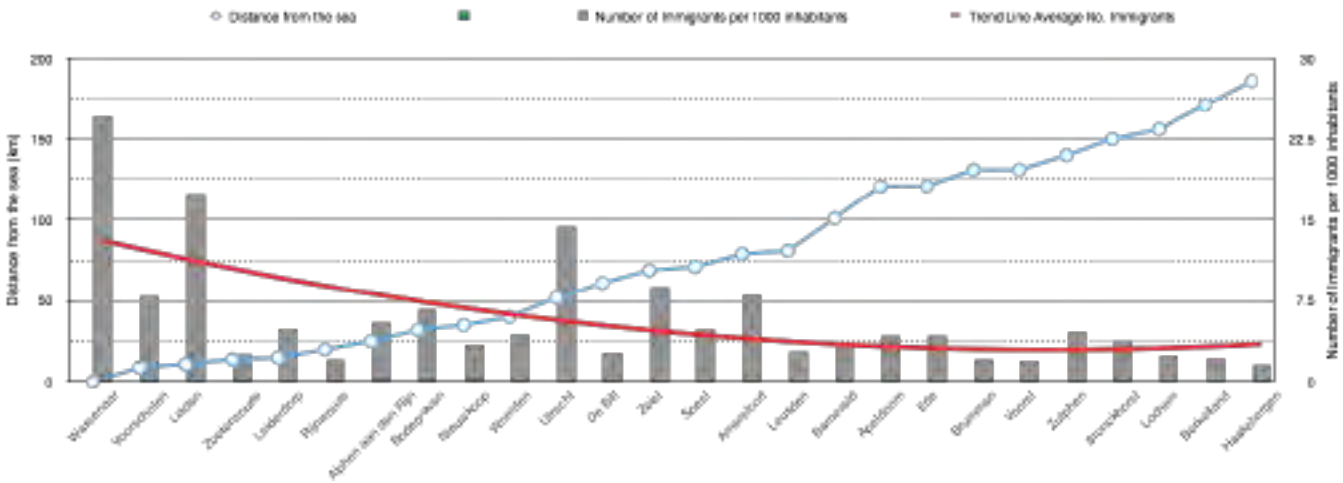


Figure 13 bar chart showing the trend between distance from the sea and no. of immigrants per 1000 inhabitants

Figure 14 a choropleth map showing the changes in number of immigrants across the transect

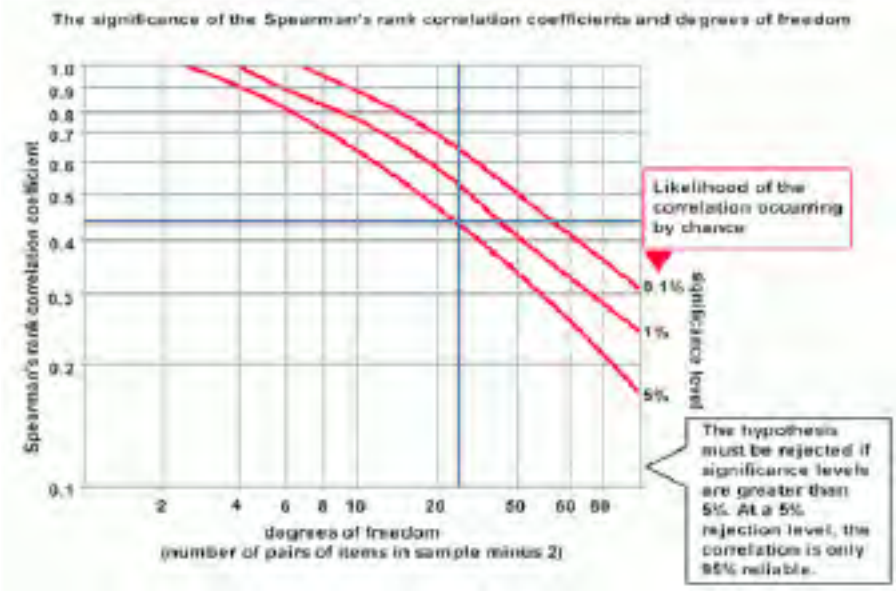
5.2 Environmental Data Results

5.2.1 Average Natural Gas Usage per Person

GM_NAAM	Natural Gas usage	rank	Distance from the sea	Rank Sea		NG difference ranks	Rank ^2
Alphen aan den Rijn	1650	3	25	7		-4	16
Amersfoort	1650	3	79	15		-12	144
Apeldoorn	3000	10	121	19		-9	81
Barneveld	2300	18	101	17		1	1
Berkelland	2300	18	171	25		-7	49
Bodegraven	2000	10	32	8		2	4
Brielselhorst	2450	24	150	23		1	1
Brummen	2300	18	131	21		-3	9
De Bilt	2400	23	61	12		11	121
Ede	2000	10	121	18		-8	64
Haakstergem	2300	18	186	26		-8	64
Leiden	1400	2	11	3		-1	1
Leiderdorp	1700	5	15	5		0	0
Leusden	2000	14	81	16		-2	4
Lochem	2500	25	156	24		1	1
Nieuwkoop	2100	15	35	9		6	36
Rijnwoude	2000	10	30	6		4	16
Soest	2200	17	71	14		3	9
Utrecht	1250	1	52	11		-10	100
Voorschoten	1900	7	9	2		5	25
Voorst	2350	22	131	20		2	4
Wenseveer	2650	26	0	1		25	625
Woerden	1950	9	40	10		-1	1
Zeist	2150	16	69	13		3	9
Zoetermeer	1900	7	14	4		3	9
Zutphen	1950	8	140	22		-14	196
						Sum D^2	1650

Table 7 (above) showing the data for natural gas usage

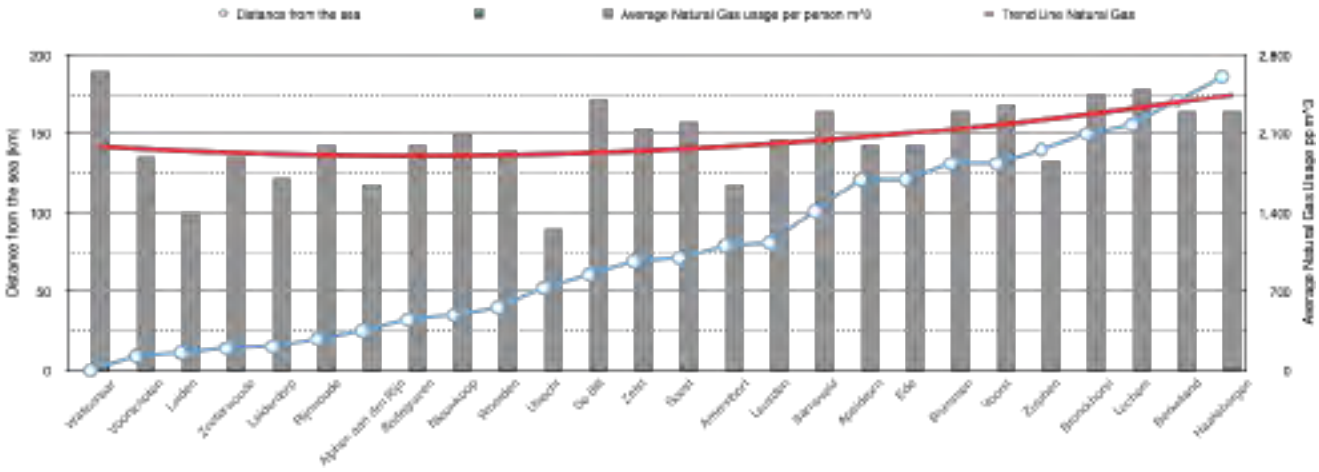
Figure 15 (below) shows the degrees of freedom graph for average natural gas usage



The Spearman's rank correlation test for average natural gas usage per person and distance from the sea produced a value of **R=0.436** which suggests that there is a positive correlation

between the two variables; when plotted on the degrees of freedom, it can be deduced that the correlation is 95% reliable, and therefore significant.

Figure 16 bar chart showing the trend between distance from the sea and average natural gas usage



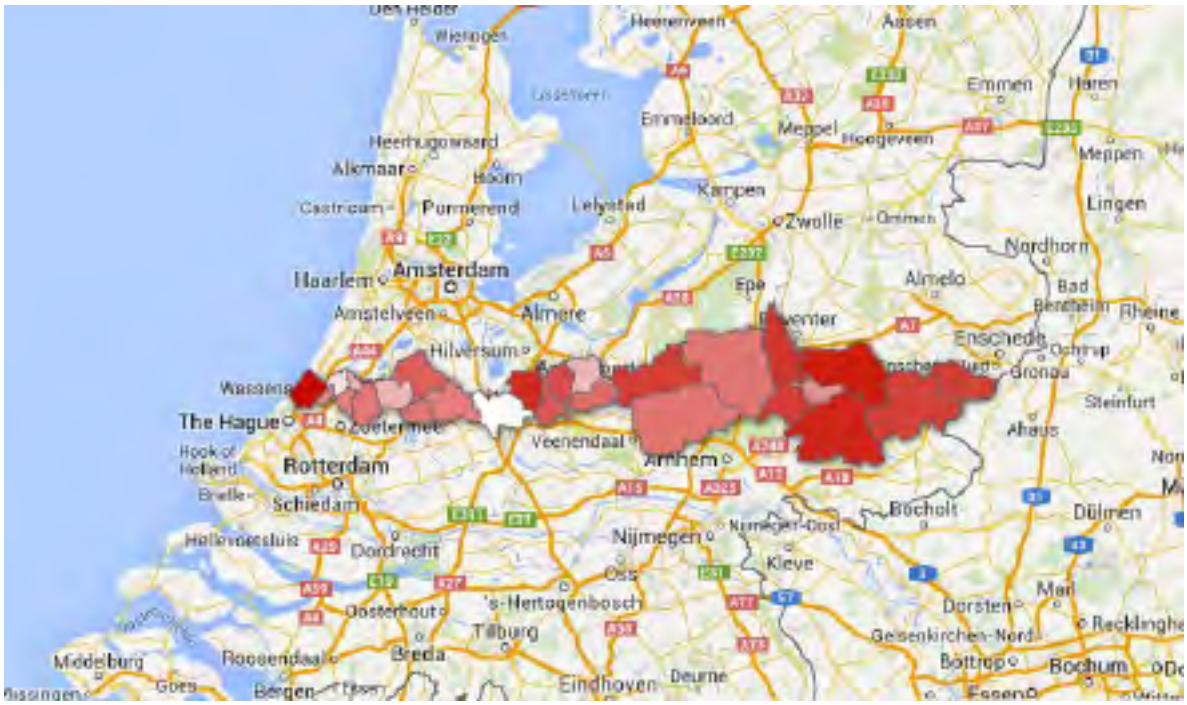


Figure 17 Choropleth map showing the changes across the transect regarding natural gas usage

The correlation shows that the average use of natural gas increases with distance from the sea as the distance from the sea is a proxy for affluence. As can be seen from Figure 16 and 17, the average natural gas usage per person varies from East Netherlands to West Netherlands. The cities in the West such as Wassenaar and Voorschoten use on average 2280 m^3 per person, whilst cities in central Netherlands use on average 1800 m^3. Lastly, the cities that are most East use on average, 2450 m^3. See Kuznets Curve #1 (figure 33).

5.2.2 Average Delivery of Electricity per Person (KwH)

GM_NAAM	Average Delivery of Electricity	Rank	Distance from the sea	Rank Sea		Electricity Diff. Rank	Electra Diff.
						0	0
Alphen aan den Rijn	2477.28	4	25	7		3	9
Amersfoort	2670.47	10	79	15		5	25
Apeldoorn	3921.04	22	121	19		3	9
Barneveld	3479.92	20	101	17		3	9
Berkelland	3436.72	19	171	25		6	36
Bodegraven	3837.61	20	32	8		7	49
Bronckhorst	3162.75	15	150	23		-12	144
Brummen	3279.56	20	131	21		10	100
De Bilt	2238.40	3	61	12		-4	16
Ede	3338.39	16	121	19		9	81
Haaksbergen	3695.79	19	186	26		1	1
Leiden	3250.23	14	11	3		7	49
Leiderdorp	2175.36	2	19	5		-11	121
Leusden	2574.31	9	81	16		-3	9
Lochem	4569.82	24	156	24		8	64
Nieuwkoop	3259.07	15	35	9		0	0
Rijnwaude	2568.90	9	20	6		6	36
Soest	2648.20	9	71	14		1	1
Utrecht	3672.30	21	52	13		-5	25
Voorschoten	1323.71	1	9	2		-10	100
Voorst	2567.26	9	131	20		-1	1
Wassenaar	2570.62	7	0	1		14	196
Woerden	2980.37	11	40	10		6	36
Zeist	3125.72	12	69	11		-1	1
Zoeterwoude	7272.23	26	14	4		-1	1
Zutphen	3387.37	17	150	22		-22	484
						5	25
					Sum Difference ^2		1622
							0.445470685408

Table 8 showing the data for average electricity delivery

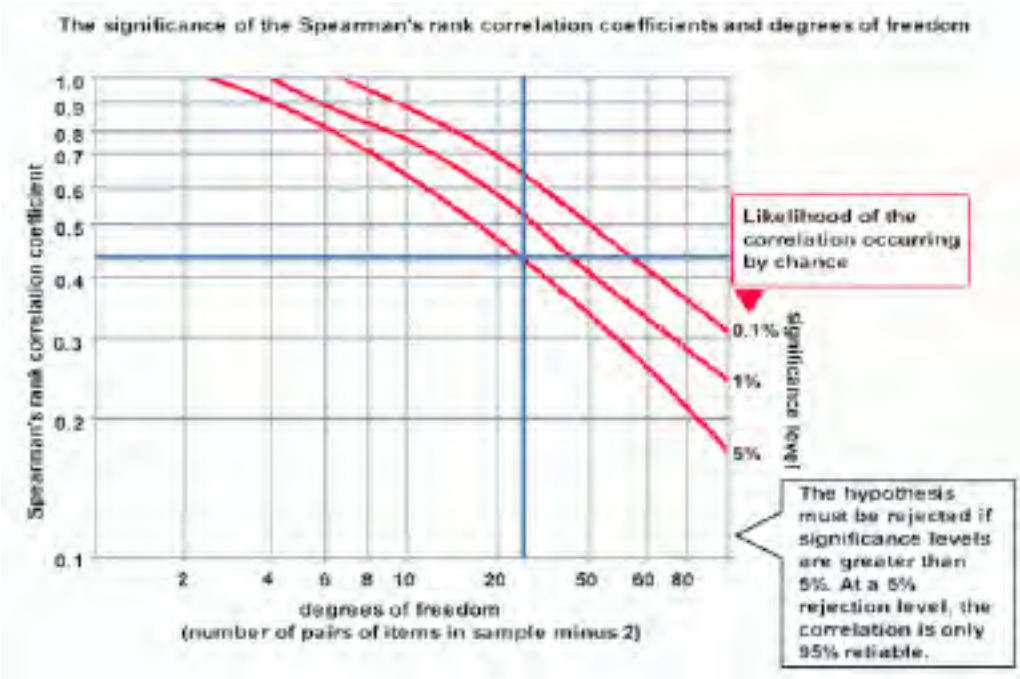


Figure 18 the degrees of freedom chart for average delivery of electricity

The spearman's rank correlation test for average delivery of electricity per person and distance from the sea produced a value of **R=0.445** and this means that there is a weak positive correlation between the two variables. When plotted on the degrees of freedom graph, the correlation is 95% reliable and therefore significant.

The correlation illustrates that the further one goes from the sea, the average delivery of electricity per person increases.

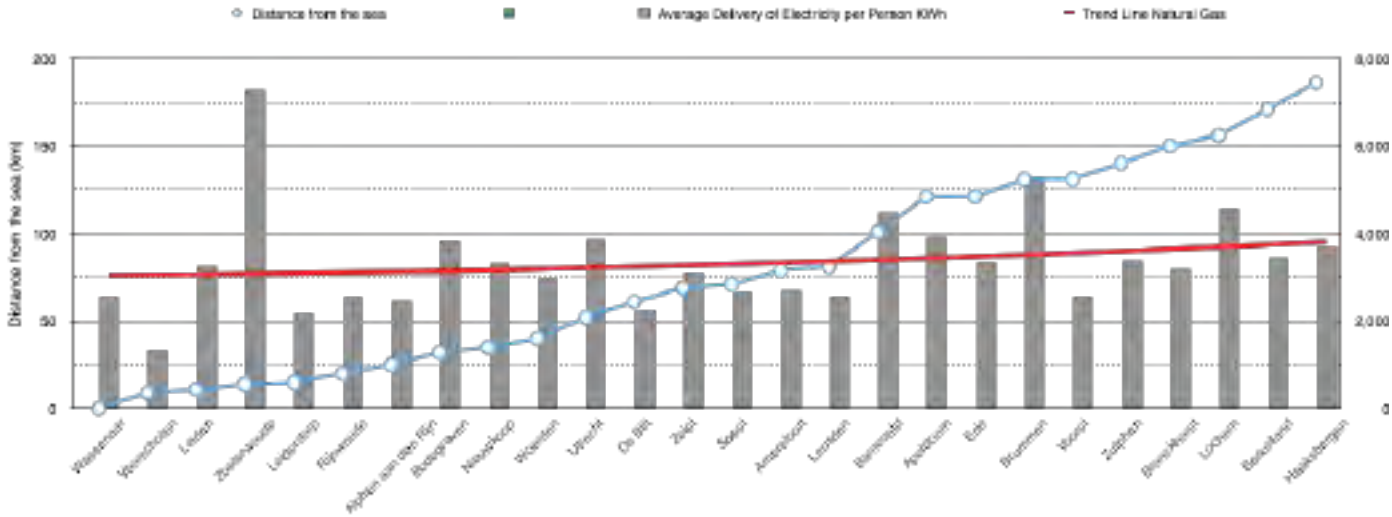


Figure 19 a bar chart showing the trend between the delivery of electricity per person and distance from the sea

Figure 20 a choropleth map showing the changes of the delivery of electricity along my transect

From the Choropleth map and graph above, it can be deduced that the trend for average delivery of electricity per person KwH is the highest in the East (less affluent). Cities on the East of The Netherlands use on average 3950 KwH of electricity, whereas central towns use 3200 KwH. Lastly, towns in the West of the country (most affluent) use on average 3000 KwH. See Kuznets curve #2 (figure 34).

5.2.3 Average Electricity Usage per Person KwH

GM_NAAM	Average Electricity Usage kWh	Rank	Distance from the sea	Rank Sea	Energy Use D	Energy Use D^2
Alphen aan den Rijn	3300	5	25	7	2	4
Amersfoort	3200	4	79	15	-11	121
Apeldoorn	3350	8	121	19	11	121
Barneveld	3650	23	101	17	-6	36
Berkelland	3650	16	171	25	9	81
Bodegraven	3500	11	32	8	-3	9
Bronckhorst	3900	26	150	23	-3	9
Brummen	3650	16	131	21	5	25
De Bilt	3650	16	61	12	4	16
Ede	3500	11	121	18	7	49
Haaksbergen	3850	25	186	26	1	1
Leiden	2850	2	11	3	1	1
Leiderdorp	3300	5	15	5	0	0
Leusden	3750	21	81	16	5	25
Lochem	3700	20	156	24	4	16
Nieuwkoop	3800	22	35	9	-13	169
Rijnwoude	3600	13	20	6	7	49
Soest	3600	13	71	14	1	1
Utrecht	2800	1	62	11	10	100
Voorschoten	3300	5	9	2	-3	9
Voorst	3650	23	131	20	-3	9
Wassenaar	3600	13	0	1	-12	144
Woerden	3450	10	40	10	0	0
Zeist	3350	8	69	13	5	25
Zoeterwoude	3650	16	14	4	-12	144
Zutphen	3650	3	140	22	19	361
					Sum Differences^2	1525

Table 9 showing the data for average electricity usage per person

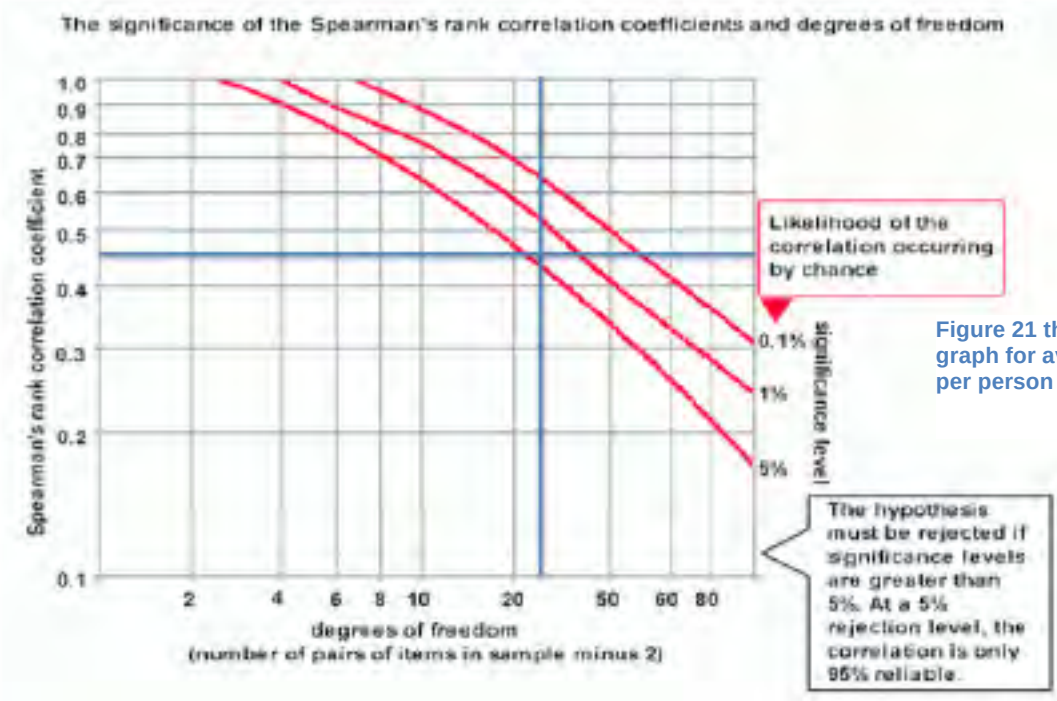


Figure 21 the degrees of freedom graph for average electicity usag per person

The spearman's rank correlation test for average electricity usage per person and distance from the sea produced a value of **R=0.479** and this means that there is a weak positive

correlation between the two variables. When plotted on the degrees of freedom graph it can seen that the correlation is more than 95% reliable, and therefore significant. The correlation shows that the average delivery of electricity per person increases, with distance from the sea.

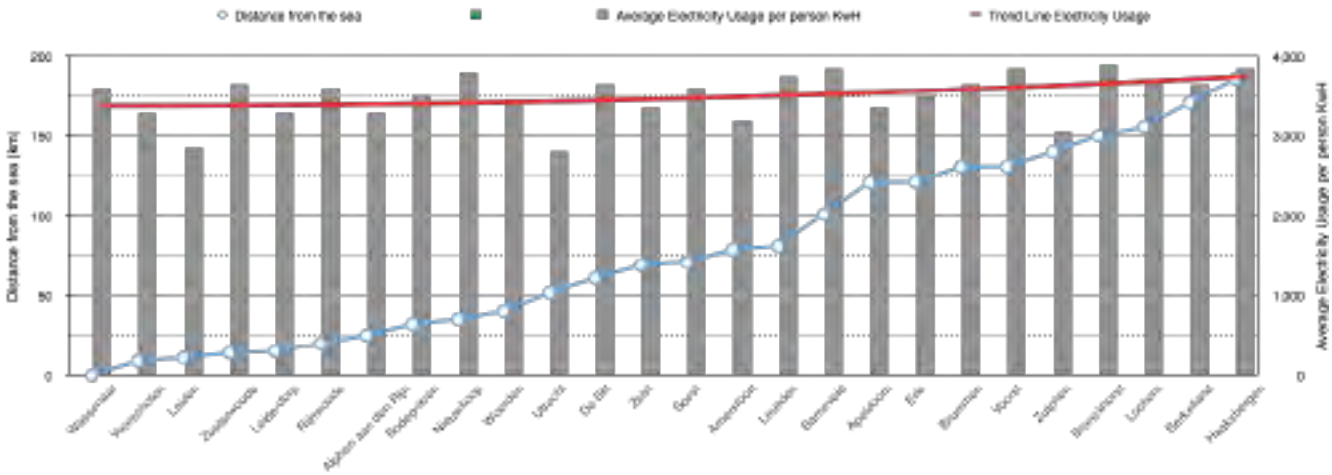


Figure 22 bar chart showing the trend of average electricity usage and distance from the sea

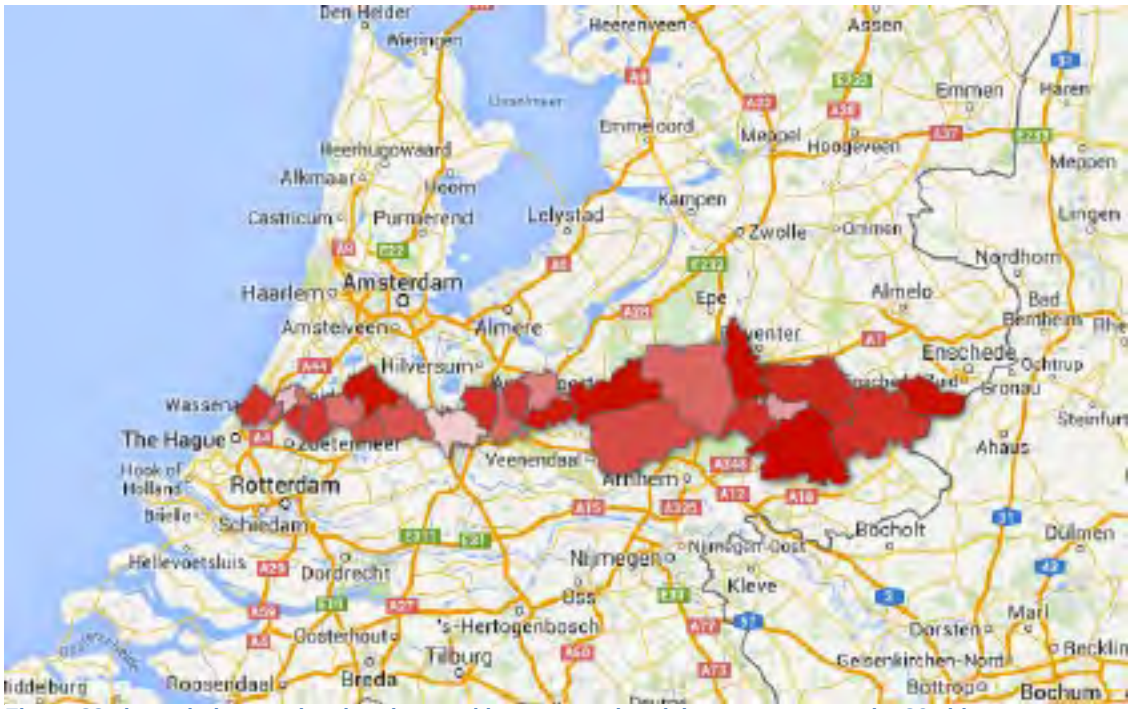


Figure 23 choropleth map showing the trend in average electricity usage across the 26 cities

The trend line of Figure 23 showing distance from the sea, a proxy for affluence, and average electricity usage per person KwH, shows that the further East you go, the average electricity usage increases from an average of 3400 kWh in the West to 3750 kWh in the East. The trend line also shows that the electricity use plateaus after a certain level affluence, however, if the trend line is extrapolated beyond Wassenaar - showing an even greater increase in affluence - the trend begins to increase again (Figure 41).
See Kuznets Curve #3 (figure 35).

5.2.4 Number of Motorbikes per 1000 people

GM_CODE	GM_NAAM	Motorbikes	Rank Motorbike	Distance from the sea	Rank Sea		Motor Bike Differenc	MB D^2
GM0484	Alphen aan den Rijn	51	13	25	7		-6	36
GM0307	Amersfoort	62	16	79	15		-3	9
GM0200	Apeldoorn	70	21	121	19		-2	4
GM0203	Barneveld	49	10	101	17		7	49
GM1859	Berkelland	40	4	171	25		21	441
GM0497	Bodegraven	34	1	32	8		7	49
GM1876	Bronckhorst	56	16	150	23		7	49
GM0213	Brummen	50	11	131	21		10	100
GM0310	De Bilt	47	9	61	12		3	9
GM0228	Ede	57	15	121	18		3	9
GM0158	Haaksbergen	39	3	186	26		23	529
GM0546	Leiden	79	24	11	3		-21	441
GM0547	Leiderdorp	81	25	15	5		-20	400
GM0327	Leusden	44	5	81	16		8	64
GM0262	Lochem	35	2	156	24		22	484
GM0569	Nieuwkoop	74	22	35	9		-13	169
GM1672	Rijnwoude	60	17	20	6		-11	121
GM0342	Soest	53	14	71	14		0	0
GM0344	Utrecht	42	6	52	11		5	25
GM0626	Voorschoten	66	19	9	2		-17	289
GM0285	Voorst	68	20	131	20		0	0
GM0629	Wassenaar	78	23	0	1		22	484
GM0632	Woerden	40	4	40	10		6	36
GM0355	Zeist	42	6	69	13		7	49
GM0638	Zoeterwoude	109	26	14	4		-22	484
GM0301	Zutphen	50	11	140	22		11	121
							Sum Difference ^2	4451
								0.512709401709402

Table 10 Showing the data for number of motorbikes per 1000 people

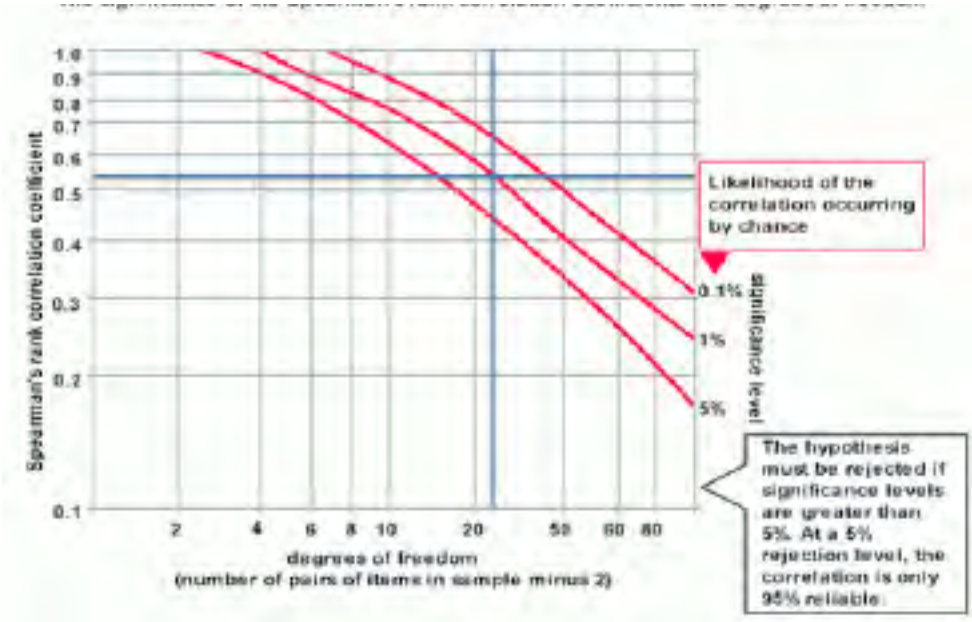


Figure 24 degrees of freedom chart for the number of motorbikes per 1000 people

The Spearman's rank correlation test for number of motorbikes per 1000 people and distance from the sea produced a value of **R=0.522** and this means that there is a strong positive correlation between the two variables. When plotted on the degrees of freedom graph we can see that the correlation is 99% reliable, and is therefore significant. The correlation shows that the further one goes from the sea, the number of motorbikes per 1000 people decreases.

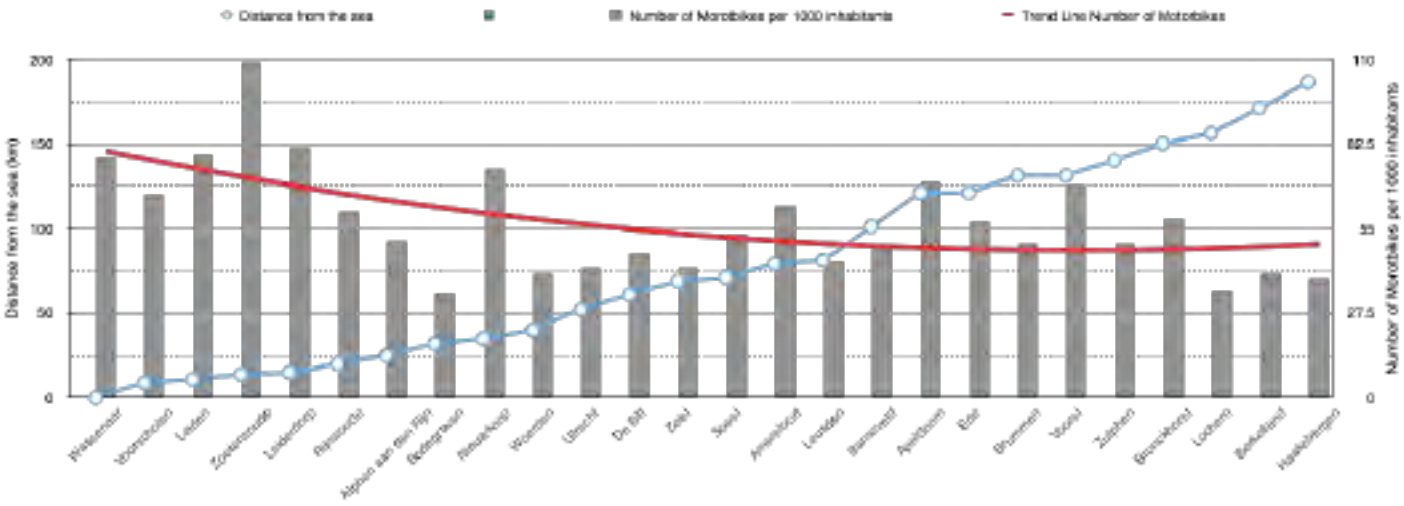


Figure 25 bar chart and trend line for the number of motorbikes and distance from the sea

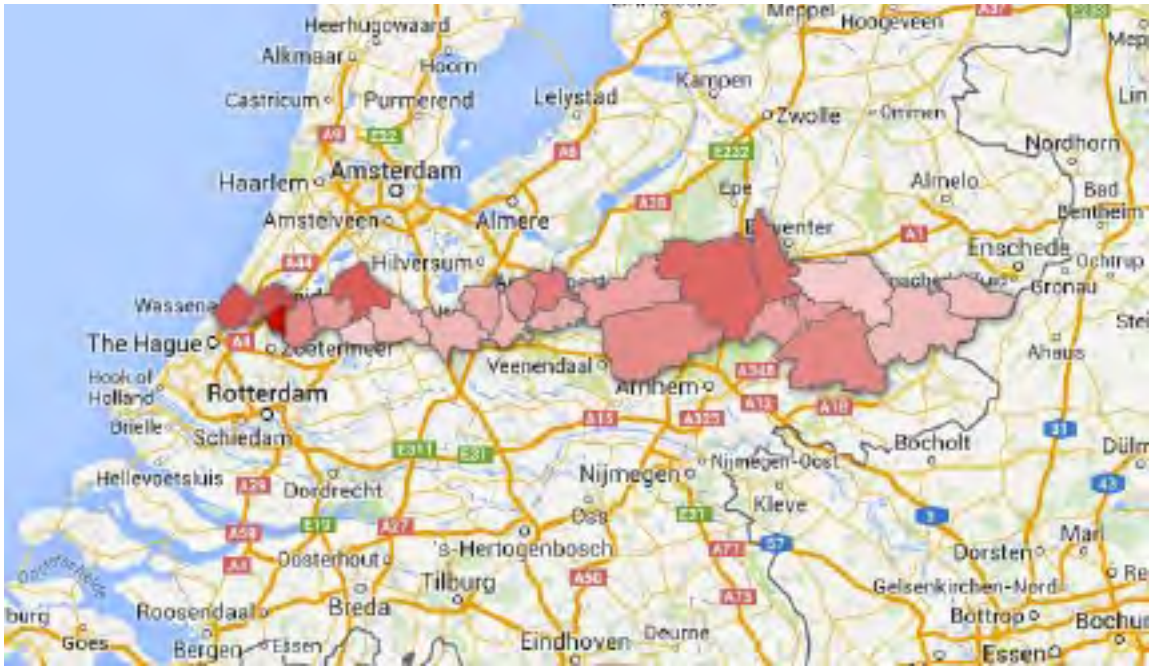


Figure 26 cloropleth map showing the changes in number of motorbikes across The Netherlands

Studying the trend of the average number of motorbikes, from the Choropleth map, it can be seen that the West of the country has a darker red and hence more motorbikes, compared to the East. When looking at the trend line, this can also be seen. However, the number of motorbikes does begin to increase again after the city of Woest, 129 km from the sea.
See Kuznets Curve #4 (figure 36).

5.2.5 Average Household Waste per KG per Person

GM_CODE	GM_NAAM	Household Waste kg per person	Rank Waste	Distance from the sea	Rank Sea		Waste D	Waste D^2
GM0484	Alphen aan den Rijn	550	11	25	7		-4	16
GM0307	Amerstoort	578	18	79	15		3	9
GM0200	Apeldoorn	468	4	121	19		-15	225
GM0203	Barneveld	576	16	101	17		-1	1
GM1659	Berkelland	577	17	171	25		-8	64
GM0497	Bodegraven	604	19	32	8		11	121
GM1876	Brenckhorst	554	13	150	23		-10	100
GM0213	Brummen	443	2	131	21		-19	361
GM0310	De Bilt	662	24	61	12		12	144
GM0228	Ede	583	14	121	19		-4	16
GM0158	Haaksbergen	667	26	166	26		0	0
GM0546	Leiden	499	7	11	3		4	16
GM0547	Leiderdorp	530	9	15	5		-4	16
GM0327	Leusden	634	23	61	12		7	49
GM0262	Lochem	553	12	156	24		-12	144
GM0569	Nieuwkoop	612	21	35	9		12	144
GM1672	Rijnwoude	496	6	20	6		0	0
GM0342	Soest	627	22	71	14		8	64
GM0344	Utrecht	524	8	52	11		-3	9
GM0626	Voorschoten	477	5	9	2		3	9
GM0086	Vaorst	684	25	131	20		5	25
GM0629	Wassenaar	610	20	8	1		19	361
GM0632	Woerden	543	10	40	10		0	0
GM0355	Zeist	575	15	69	13		2	4
GM0638	Zoeterwoude	368	1	14	4		-3	9
GM0301	Zutphen	484	3	140	22		-19	361
							Sum Difference ^2	2268
								0.224615384615385

Table 11 showing the data regarding average household waste per KG per person

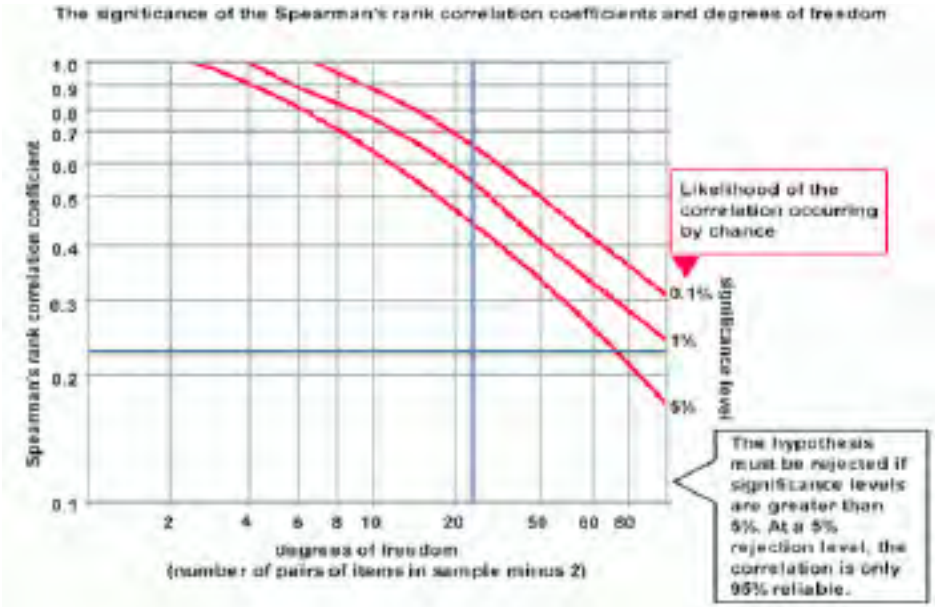


Figure 27 degrees of freedom chart for average household waste per KG per person

The spearman's rank correlation test for household waste per person KG and distance from the sea produced a value of **R=0.225** and this means that there is a very weak positive correlation between the two variables. When plotted on the degrees of freedom, it can be seen that the correlation has less than 95% reliability; as a result the correlation is not significant and cannot be used further.

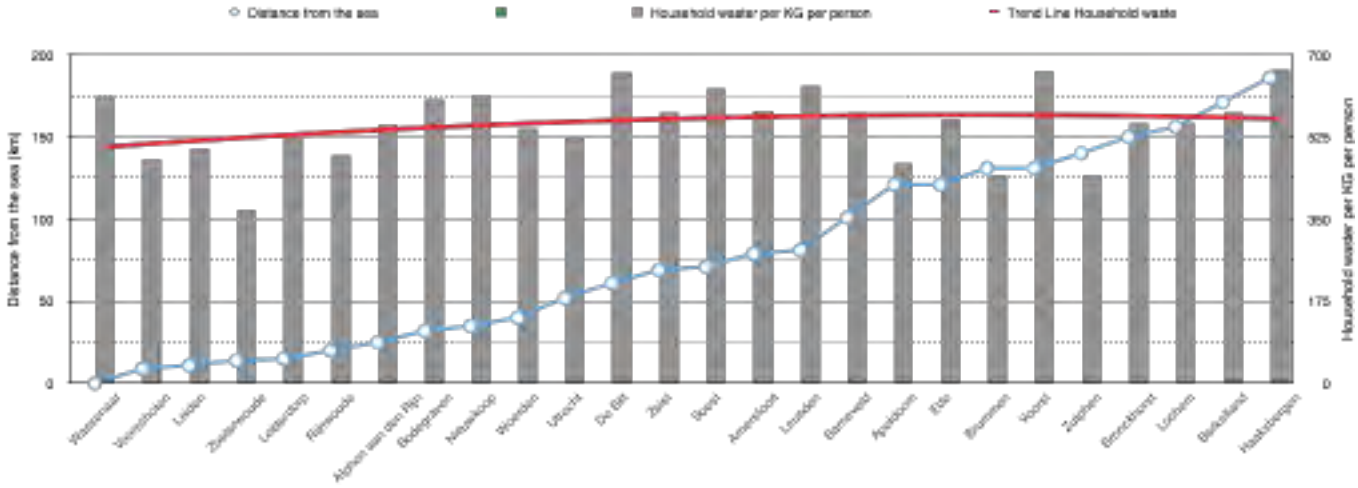


Figure 28 trend-line showing the changes in household waste and distance from the sea (affluence)

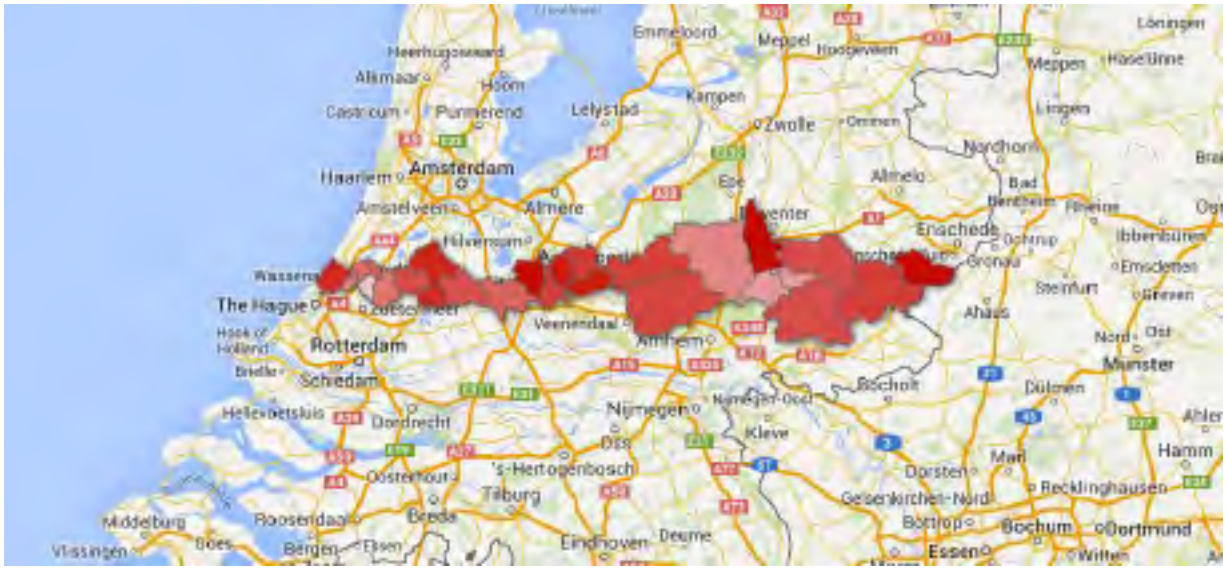


Figure 29 cloropleth map showing the changes in household waste along the transect

Household waste per KG per person is peculiar, as the trend line has not followed the expected trend lines for the other environmental factors. For household waste, the central region of The Netherlands has the highest value for household waste, whilst the eastern and western most cities have the lower values for household waste.

See Kuznets Curve #5 (figure 37). The Kuznets curve for this environmental factor was drawn and used as a comparison in my analysis, however, since the correlation is not significant, we cannot use this factor in our conclusion.

5.2.6 Average Number of Cars Per Person

Table 12 showing the data for the number of cars per person

GM_CODE	GM_NAAM	Personal Cars	Rank Cars	Distance from the sea	Rank Sea		Difference in Ranks	Difference Ranks^2
GM0484	Alphen aan den Rijn	0,97152541	17	25	7		-10	100
GM0307	Amersfoort	0,87943430	8	79	15		7	49
GM0200	Apeldoorn	0,97625656	19	121	19		0	0
GM0203	Barneveld	0,96913399	15	101	17		2	4
GM1859	Berkelland	0,62477716	6	171	25		19	361
GM0497	Bodegraven	0,41344383	2	32	8		6	36
GM1876	Bronckhorst	0,85272229	26	150	23		-3	9
GM0213	Brummen	0,50017706	3	131	21		-18	324
GM0310	De Bilt	0,08119219	1	61	12		11	121
GM0228	Ede	0,96940577	16	121	18		2	4
GM0158	Haaksbergen	0,97299105	18	186	26		8	64
GM0546	Leiden	0,99055477	23	11	3		-20	400
GM0547	Leiderdorp	0,98182268	21	15	5		-16	256
GM0327	Leusden	0,96383948	13	81	16		3	9
GM0262	Lochem	0,63868147	7	156	24		17	289
GM0569	Nieuwkoop	0,57674912	5	35	9		4	16
GM1672	Rijnwaarde	0,55474917	4	20	6		2	4
GM0342	Soest	0,97786211	20	71	14		-6	36
GM0344	Utrecht	0,94191103	11	52	11		0	0
GM0626	Voorschoten	0,96166215	10	9	2		-8	64
GM0285	Voorst	0,96201127	12	131	20		8	64
GM0629	Wassenaar	0,04003951	25	0	1		-24	576
GM0632	Woerden	0,95458649	9	40	10		1	1
GM0355	Zeist	0,99740771	24	69	13		-11	121
GM0638	Zoetermeer	0,98557823	22	14	4		-18	324
GM0301	Zutphen	0,96662913	14	140	22		8	64
						Sum Difference ^2		3246

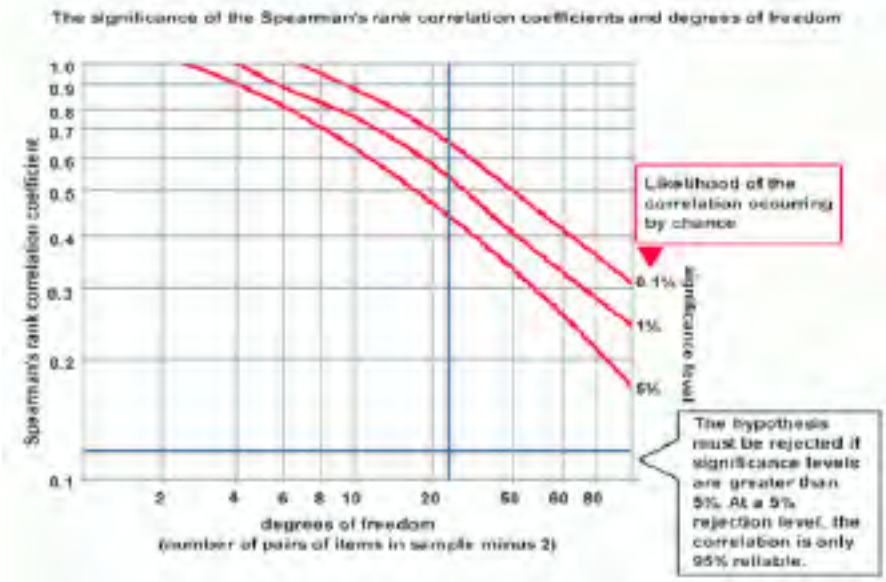


Figure 30 showing the degrees of freedom chart for the number of cars per person

Table 12 shows the data for number of cars per person, and distance from the sea as well as their ranks. The last two columns (in yellow) show the Spearman's rank correlation coefficient calculations for this environmental factor.

The spearman's rank calculation produced a value of **R=-0.127, which suggests a very weak negative correlation.** When plotted on the degree of freedom graph it produced less than 95% reliability. As a result, this correlation is not significant and must be rejected.

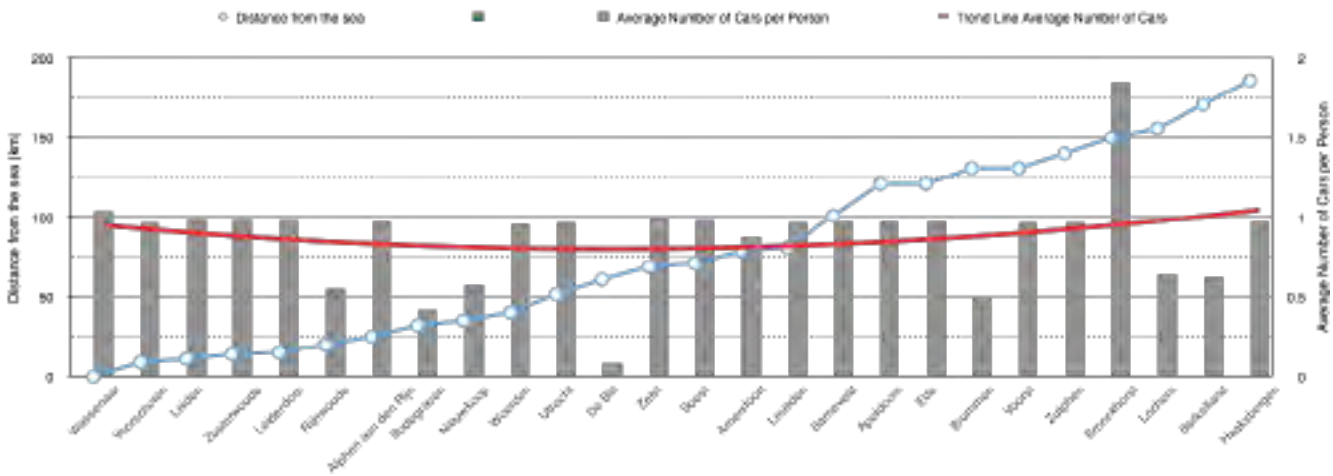
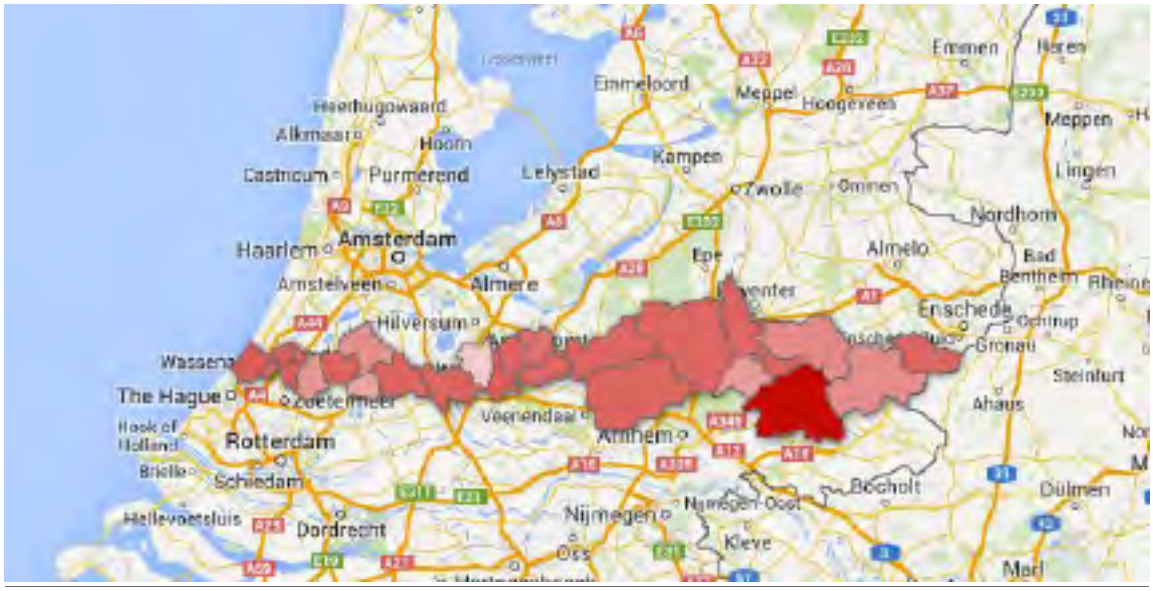


Figure 31 bar chart with trend line showing changes number of cars per person

Figure 32 a choropleth map showing the changes in the number of cars per person



This environmental factor of number of cars per person, even though having a fairly low correlation, has a distinct trend line showing the central region of the country having the lowest number of cars per person (hence less pollution) whilst the eastern and western most cities have the highest number of cars. This trendline for personal cars matches the expected trend line of the other environmental factors of The Netherlands.
See Kuznets Curve #6 (figure 38). The Kuznets curve for this environmental factor was drawn and used solely for comparison purposes in my analysis, however, since the correlation is not significant, we cannot use this factor in our final conclusion.

6 Analysis

6.1 Kuznets Curve's

Figure 33 The Kuznets Curve for average natural gas usage

Figure 34 The Kuznets Curve for average delivery of electricity per person

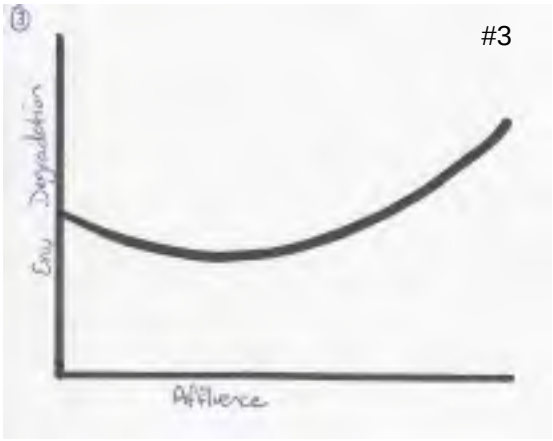


Figure 35 The Kuznets Curve for average electiricty usage

Figure 36 The Kuznets Curve for ave. number of motorbikes per person

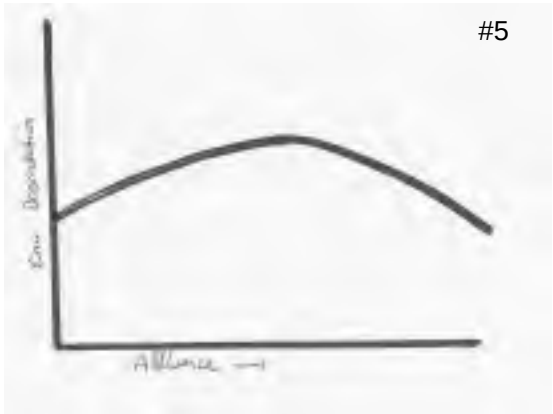


Figure 37 The Kuznets Curve for ave. household waste per person

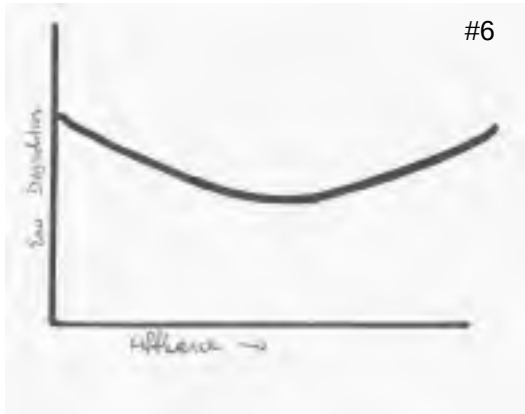


Figure 38 The Kuznets Curve for number of cars per person

The Kuznets curves shown above for each of the 6 environmental factors, the general trend for the curve in The Netherlands is the inverse of the published Kuznets Curve (figure 42 + 43). If this inverse shape of Kuznets curve for The Netherlands is true, then using my 'Ideal City' calculations (**see chapter 4.2),** the most environmentally friendly city (the one with the lowest environmental degradation should be in the middle of the transect of The Netherlands.

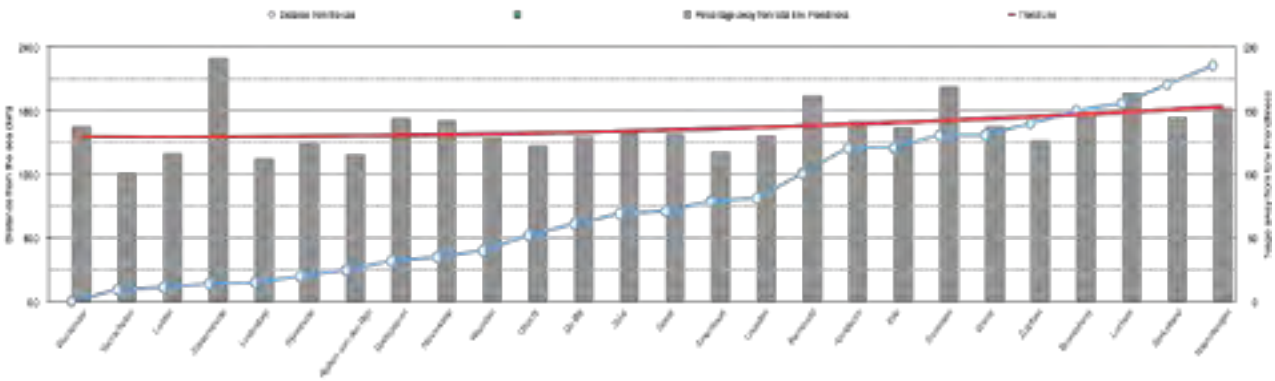
From the table on the left showing the percentage a city is away from being as environmentally friendly, it can be seen that the city 'De Bilt' is just 1.67% from being the most sustainable city. De Bilt is the twelfth city from the coast in my transect of 26 cities, and when plotted on the map, 'De Bilt' is situated in the middle of The Netherlands, hence following the predicted theory.

	Final Percent	Rank Sea
Wassenaar	165.234478026243	1
Voorschoten	123.118665476202	2
Leiden	106.230699855004	3
Zoeterwoude	120.233363766791	4
Leiderdorp	114.28775335668	5
Rijnwoude	108.411805104538	6
Alphen aan den	108.753902560085	7
Bodegraven	130.668808696196	8
Nieuwkoop	109.978214136205	9
Woerden	114.520369832381	10
Utrecht	107.325479261642	11
De Bilt	101.667933809902	12
Zeist	121.711830739654	13
Soest	116.551987611474	14
Amersfoort	111.291811308239	15
Leusden	117.95594065487	16
Barneveld	107.744323319971	17
Apeldoorn	106.305175753009	18
Ede	105.065936209852	18
Brummen	106.038540308708	20
Bronckhorst	110.624939054676	23
Lochem	113.934378585178	24
Berkelland	98.4575147235475	25
Haaksbergen	102.620524457471	26

Table 13 a table showing the least environmentally degrading city

Figure 39 a map showing the location of the least environmentally degrading city

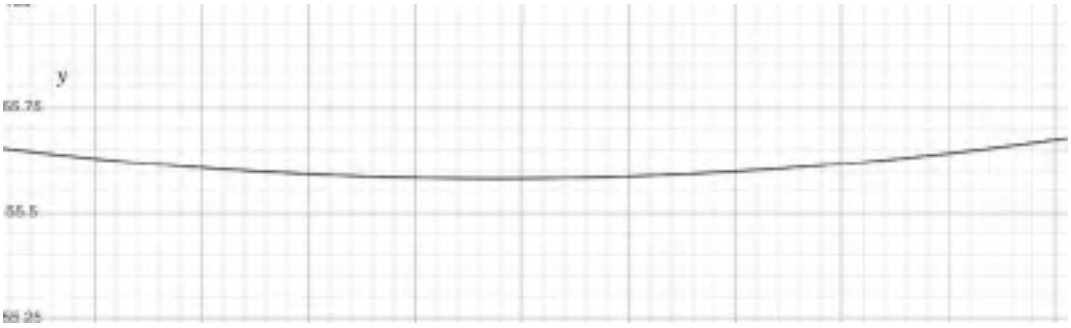
Figure 40 a bar chart and trend lines to show the changes in environmental degradation across the 26 cities



When plotting the environmental data alone, the percentage a city is away from 100% (showing total environmental friendliness) shows how environmentally degrading the region is.

The trend line clearly shows, that the furthest cities from the sea (signifying the least affluent) have a higher level of environmental degradation compared to the middle-income cities. The cities closest to the sea (therefore, by proxy, the most affluent), have an equal environmental degradation score compared to the cities in the middle of the transect, however, the trend line does begin to rise again towards the most affluent three cities. When the trend line was extrapolated, it was concluded that the level environmental degradation began to increase again.

Figure 41 the trend line for the perfect city extrapolated



The Kuznets patterns drawn for The Netherlands, seems to be reversed, however there may be some economic factors that have caused this.

Figure 42 The published Kuznets Curve of Env. Degradation

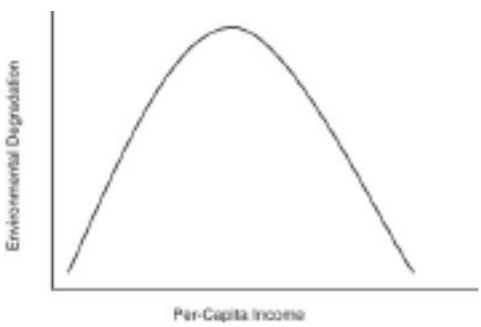


Figure 43 The Kuznets Curve of Env. degradation applicable for The Netherlands

7 Discussion

7.1 Market Forces

When researching possible reasons for why the poorer regions in The Netherlands, the cities most East, have an increased environmental degradation when compared to the richest cities in The Netherlands, it was found that the economic requirements such as electricity prices and unemployment benefits varied across the transect.

When comparing the same variables of 5100 KwH per year of electricity, and 1800 m^2 of natural gas per year, the average price for the energy in Wassenaar was around €4073 - Wassenaar marks the richest city. Furthermore, middle-income cities, ie. Leusden located in the heart of The Netherlands was analyzed with the same variables; the average energy price in Leusden was €2443. Lastly, the least affluent and furthest East city was evaluated; Haaksbergen's energy prices were just under €2200. When comparing the decreasing prices in energy with the country's economic data, (figure 10) as well as the average household electricity usage and natural gas usage - figure 23 and 15 respectively, it can be seen that the least affluent cities have an increased usage of both energy types than any other city on the transect. This increase, could be due to the reduction in energy prices and hence, the citizens in these cities are able to use a greater amount energy for a lower fee.

In addition, the Dutch government pays unemployed workers 70% of their previous salary (I Amsterdam 2014). The percentage of the cities population of who are claiming benefits was plotted with the affluence of the area (Figure 44).

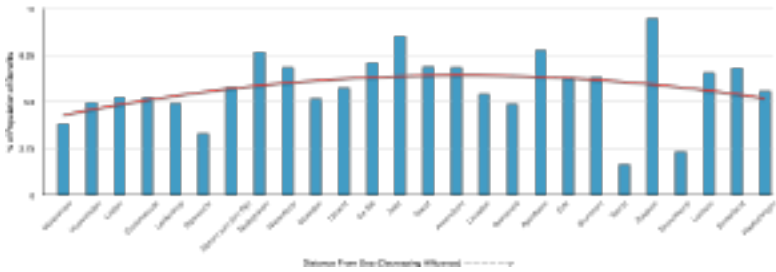


Figure 44 Unemployment across the 26 cities

Figure 44 shows an increase in benefits for the middle-income regions such as De Bilt and Apeldoorn, whereas the poorest and richest regions have a decrease in unemployment benefits. Research done in America (figure 45) looking at the correlation between employment and energy consumption shows that

'Economic growth is tied to job creation, so it stands to reason that energy consumption would be tied to job growth. But I will have to admit that I was surprised by the closeness of the relationship for the period shown.' (Tverberg 2012).

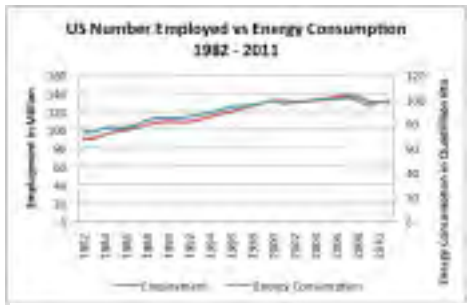


Figure 45 a graph showing the US number of employed peoples and energy consumption

With the knowledge gained from the correlation of employment and energy consumption, the cities along the transect with the lowest unemployment are the

With all the market forces and the environmental data combined, the poorest regions in The Netherlands, along with the higher employment and lower energy prices, consume the same amount, if not more, than the most affluent areas in The Netherlands. Hence, they have a higher-level environmental degradation than the middle-income cities. The Kuznets Curve found for the transect going is shown below (figure 46).

Figure 46 environmental degradation with affluence in The Netherlands

7.2 Improvements and Limitations

When collecting data from the CBS StatLine, especially for some of the smaller cities like Leusden, the census data available was relatively out of date. As a result, it was attempted to average the data from the year 2010, as this was the most widely available year. However for one or two variables such as average electricity delivery, the latest available year, where there was data for all 26 cities, was 2008. This was post financial crisis and therefore the values could have been significantly different than those taken in 2010, for example.

Moreover, due to data restrictions, only one transect across the Netherlands was delineated, however the results could be different if another transect going East to West was chosen at a higher or lower latitude. The transect used includes some of the wealthiest and some of the poorest regions in The Netherlands, and these extremes in affluence could also lead to extreme data results; hence, when working out averages and trends, these extreme values could cause deviations, and the extremes could mask the ‘accurate’ or expected trend.

Some of the environmental data used had a Spearman's rank value of below 0.44 and for my sample size; anything below 0.44 has to be considered an invalid hypothesis. Number of cars per person and average household waste per person in kilograms had Spearman’s values of below this and therefore the correlation between these two environmental factors had to be considered invalid.

7.3 Evaluation of Hypothesis

Reverting back to the original hypothesis stated in section 3.1.

The results, trend lines and analyzed market forces, show that the original hypothesis was correct. The Kuznets Curve that has been modified for the Netherlands (figure 46) shows that the most and the least affluent areas have the highest level of environmental degradation, whilst the middle-income regions have the lowest level of environmental degradation.

8 Conclusion

In conclusion, the East-West transect in the Netherlands, of 26 cities, follow a different pattern to the published Kuznets Curve.

For each of the six environmental factors, the bar graph and trend line show the changes in the variable with distance from sea, a proxy for affluence. The trend lines for all but 1 environmental factor showed the **same** pattern: the most affluent and the least affluent regions in The Netherlands had the highest level of environmental degradation, whereas the middle income cities had the lowest. The market forces analyzed - electricity and gas prices, as well as unemployment also proved this relationship between affluence and environmental degradation. The Netherlands therefore, does fully not follow the Kuznets Curve, however, a more appropriate representation of environmental degradation in The Netherlands would be the inverse of the published Kuznets Curve (a U shape).

There are a wide variety of other market forces that can affect the environmental degradation of a region, including the efficiency of waste disposal, sustainability education, percentage of young persons compared to older people and many more. All of these will affect the mentality and environmentally friendliness of the city.

The future may hold changes in the mentality of Dutch citizens, and it will be interesting to look at the changes in the 26 cities in the future as the economy redevelops and compare these changes to data from post financial crisis (2008).

9 Bibliography

"The Netherlands." *Central Intelligence Agency*. Central Intelligence Agency, n.d. Web. 25 June 2014. <https://www.cia.gov/library/publications/the-world-factbook/geos/nl.html>.

"Timeline Dutch History." *History of The Netherlands*. De Rijksmuseum, n.d. Web. 25 June 2014. <https://www.rijksmuseum.nl/en/explore-the-collection/timeline-dutch-history>.

Bijlo, Esther. "Ongelijkheid Tussen Arm En Rijk Verrassend Groot in Nederland." *De Verdieping*. Trouw, 05 June 2014. Web. 25 June 2014. <http://www.trouw.nl/tr/nl/4504/Economie/article/detail/3667222/2014/06/05/Ongelijkheid-tussen-arm-en-rijk-verrassend-groot-in-Nederland.dhtml>.

Bregman, Rutger. "Hoe Groot Is De Ongelijkheid in Nederland?" *Hoe Groot Is De Ongelijkheid in Nederland?* De Correspondent, Oct. 2013. Web. 25 June 2014. <https://decorrespondent.nl/55/hoe-groot-is-de-ongelijkheid-in-nederland/2114475-04df073f>.

"What Is Gini Index? Definition and Meaning." *BusinessDictionary.com*. N.p., n.d. Web. 25 June 2014. <http://www.businessdictionary.com/definition/gini-index.html>.

Salverda, W, Haas, C, De Graaf-Zijl, M, Lancee, B, Notten, N, and Ooms, T. "Growing Inequalities and Their Impacts in The Netherlands." GINI Growing Inequalities' Impacts, Sept. 2013. Web. 25 June 2014. <http://www.gini-research.org/system/uploads/512/original/Netherlands.pdf?1380138293>.

"Unemployment benefits in the Netherlands." *I Amsterdam*. N.p., n.d. Web. 19 July 2014. <http://www.iamsterdam.com/en-GB/living/Work/employment-law-and-benefits/unemployment-benefits>.

Tverberg, Gail. "The Close Tie Between Energy Consumption, Employment, and Recession." *Our Finite World* 1 (2012): n. pag. *Our Finite World*. Web. 19 July 2014. <http://ourfiniteworld.com/2012/09/17/the-close-tie-between-energy-consumption-employment-and-recession>.

Stern, David I. "The Environmental Kuznets Curve." *International Society for Ecological Economics* 1 (2003): n. pag.*The International Society for Ecological Economics*. Web. 24 July 2014. <http://isecoeco.org/pdf/stern.pdf>

Anderson, Michael W. "The Dutch Economy: A History of the Dutch Economy since WWII." *Michael W Andersons Blog*. N.p., 2 Jan. 2010. Web. 24 July 2014 <http://andersonmwa.wordpress.com/2010/02/01/the-dutch-economy-a-history-of-the-dutch-economy-since-wwii-2/>.

"Countries Compared by Geography." *NationMaster*. NationMaster, n.d. Web. 31 July 2014. <http://www.nationmaster.com/country-info/stats/Geography/Area/Comparative>.

Investigating factors affecting presence of *Propithecus coquereli* in the dry deciduous forest of Mariarano in North-West Madagascar, and the implications for the conservation of this endangered species

To what extent are Coquerel's sifakas *(Propithecus coquereli)* selective in their tree use based on tree height in the dry deciduous forest of Mariarano, Northwest Madagascar? And further, to what extent are *P. coquereli* selective in their tree use based on species of tree?

Ale Baranowski - Biology

When I starting considering topics for my Extended Essay, researching the behaviour of Coquerel's Sifakas in Madagascar could not have been further from my mind. But, taking our trip to Madagascar to explore this field of Biology previously so unfamiliar to me proved to be an unexpectedly enlightening decision. Working with the scientists at the conservation camps, I gained an insight into their work for the endangered species and cultivated my own understanding of various methods of analysis and areas of Ecology that we could not have covered within the syllabus. The process of completing my Extended Essay widened my knowledge and appreciation of the significance of conservation in one of the world's most biodiverse areas, and allowed me to work independently when collecting the data and analysing various aspects of it. Most importantly, however, it has given me the chance to see what original research actually entails, through the moments of enlightenment but also the countless setbacks, thus giving me a taste of what university research might entail.

Supervisor: Hélène Bonsall

Once Alexandra set her mind on the study of lemurs in Madagascar she immersed herself in her project passionately and without further hesitation, even in the face of challenging circumstances such as the long, dusty walks to collect her data and overcoming a language barrier to be able to communicate with the researchers in the area. I was impressed with her intellectual initiative as she carried out her research independently, particularly with regards to finding out about statistical methods she could use to analyse her data. Despite the fact that these statistical skills and the fieldwork techniques were completely new to her, she showed genuine biological flair by applying them as effectively as she did.

Contents

Abstract

This investigation sought to answer two questions concerning selective habitat use by Coquerel's sifakas. These read "To what extent are Coquerel's sifakas *(Propithecus coquereli)* selective in their tree use based on tree height in the dry deciduous forest of Mariarano, Northwest Madagascar? And further, to what extent are *P. coquereli* selective in their tree use based on species of tree?" It was found that there is a significant difference in the average tree height of randomly selected trees and those selected by *P. coquereli*, supporting the initial hypothesis, which predicted the sifakas preferring taller (>15m) trees. Results for the effect of tree species were less powerful, not showing a significant preference by *P. coquereli*. However, the data suggests that some selective behaviour is shown towards species of tree. Further research could enhance these findings. Fieldwork was carried out over six weeks in the Malagasy dry season in Mariarano, Northwest Madagascar, the area to which *P. coquereli* are endemic. Preliminary work was conducted in the form of opportunistic surveys in the five routes surrounding the Mariarano base camp, and subsequently line transect sampling was carried out, with each route being walked a minimum of 6 times throughout the course of the investigation. A group of three observers were escorted by a local guide, walking the transects and scanning for groups of *P. coquereli.* When a sighting occurred, the PCQ method was utilised, providing information for the central tree and four quarter trees surrounding it. This study is aimed at enriching data available on *P. coquereli* in order to enhance conservation efforts and strengthen the case presented to NGOs and the Malagasy Government for expanding protected areas in the Boeny area. It also seeks to lead to an effective reforestation effort to ensure the ability of this endangered species to survive and thrive in an artificially replanted habitat.

INTRODUCTION

1.1 RESEARCH QUESTIONS

To what extent are Coquerel's sifakas (*Propithecus coquereli)* selective in their tree use based on tree height in the dry deciduous forest of Mariarano, Northwest Madagascar? And further, to what extent are *P. coquereli* selective in their tree use based on species of tree?

1.2 BIO-GEOGRAPHICAL BACKGROUND OF MADAGASCAR AND ITS SPECIES

Ever since watching the film *Madagascar* as a child, the image of King Julian stayed with me, and so the thought of seeing and working with but more importantly working *for* these animals had remained in the back of my mind, in the hopes of one day being fulfilled. Hearing of the opportunity to visit Madagascar over summer to work alongside "Operation Wallacea", I now returned to these childhood aspirations, but this time from a different perspective - that of aiding the conservation efforts of the scientists working incessantly to create sustainable ways of protecting this unique environment and the immense wealth of fauna and flora found within it *(see **section 1.4**).* During the first surveys I became progressively intrigued as to whether lemurs are selective in their tree use and if so, which were the primary factors causing them to do this. This led me to forming the research questions "To what extent are Coquerel's sifakas (*Propithecus coquereli)* selective in their tree use based on tree height in the dry deciduous forest of Mariarano, Northwest Madagascar? And further, to what extent are *P. coquereli* selective in their tree use based on species of tree?"

Madagascar is a biodiversity hotspot[1] and presents such diversity in its fauna and flora that it is often referred to as the "eighth continent"[2], with 75% of 115 species of mammals being endemic.[3] Since Madagascar split from mainland Africa around 130 million years ago, the isolationist evolution that followed has been the most significant cause of the high levels of endemism, which in primates surpass those of any other nation, with all 103 taxa being endemic.[4]

Varied bioclimatic zones across the island[5] have also led to micro endemism, where species' presence is restricted to certain regions. This phenomenon further exacerbates the problem caused by illegal logging and slash-and-burn activities, as species cannot migrate to other parts of the island if their habitats get destroyed or fragmented. By not being able to acclimatise and adapt to living conditions, extinction becomes a major threat.

1.3 ANTHROPOGENIC THREATS TO THE HABITATS OF *P. COQUERELI*

Anthropogenic threats are prominent and progressively increasing, with fragmentation of habitats and deforestation occurring rapidly throughout Madagascar, primarily due to slash-and-burn activities used by locals. Being one of the poorest nations, with 92% of the population living below the poverty line in 2013,[6] the rich variety of the forest's raw materials has led a dependency on them by the locals. They are used as construction materials, firewood or treatment of illnesses to name a few. Large corporations, who exploit and export rare but valuable wood such as ebony and rosewood, further pose a prominent threat. Political instability, corruption and a lack of funding have made it exceedingly difficult to monitor the nation's forests. Furthermore, though

[1] Myers (1988)
[2] Ferguson (2014)
[3] Mittermeier *et al.* (2013)
[4] Mittermeier *et al.* (2013)
[5] Dry deciduous forests in the North and West, humid climate in the East and spiny thickets in the South.
[6] http://www.worldbank.org/en/news/feature/2013/06/05/madagascar-measuring-the-impact-of-the-political-crisis

traditionally it has been taboo or *fady* to poach *P. coquereli,* the arrival of migrants who do not respect these traditions has escalated hunting activity, and poaching is now a significant threat, as is the exotic pet trade.[7]

1.4 SIGNIFICANCE OF CONSERVATION EFFORTS OF *P. COQUERELI*

Coquerel's sifakas (*Propithecus coquereli)(see app. A)* are found solely in the dry deciduous forests of North West Madagascar. These arboreal primates thrived in Madagascar in the absence of more recent primate predators such as apes, and can reach 110cm in height. They have predominantly white pelage with some brown[8], and usually live in groups of four to five individuals. Apart from promoting ecotourism in Madagascar as a flagship species, *P. coquereli* has been used in research concerning human evolutionary history and social behaviour in matriarchal societies.[9] Increasing ecotourism could increase the country's GDP significantly, even leading to funding by NGOs or foreign individuals. It would simultaneously increase awareness of both the locals and foreign travellers on the importance of conservation, and create employment opportunities for the Malagasy in protected areas and educating. Thus, the quality of their living standards could increase greatly. *P. coquereli* is listed as "Endangered" in the IUCN red list, and Schwitzer *et al.* showed in 2013 that over a period of three generations (c. 50 years) their population has halved.[10]

To aid conservation efforts, amount of data held on *P. coquereli* must be increased as very little research has focused on this species. Studying the factors leading to selective tree use by *P. coquereli* is highly significant as it will facilitate conservationists with forest replantation and habitat restoration, to recreate an environment as undifferentiated as possible from the original. As of 2010, just 8%

[7] http://news.mongabay.com/2009/0330-hance_madagascar.html
[8] http://www.arkive.org/coquerels-sifaka/propithecus-coquereli/image-G1568.html
[9] Grieser (1992)
[10] http://www.iucnredlist.org/details/18355/0

of the nation's total land area is protected,[11] and of the area inhabited by *P. coquereli,* less than 10% is protected *(see **Fig. 1,** next page).* Substantial data will aid efforts made by NGOs such as Operation Wallacea and DBCAM to lobby for the increase of protected areas within the Mahamavo region. It will also strengthen the efforts to attract desperately needed foreign aid and funding, much of which was cut off with the rise of the political instability in 2009.[12]

Research was to be carried out over six weeks, from June 21st to August 1st 2014 during the Malagasy dry season.

[11] Mittermeier *et al.* (2013)
[12] Mittermeier *et al.* (2013)

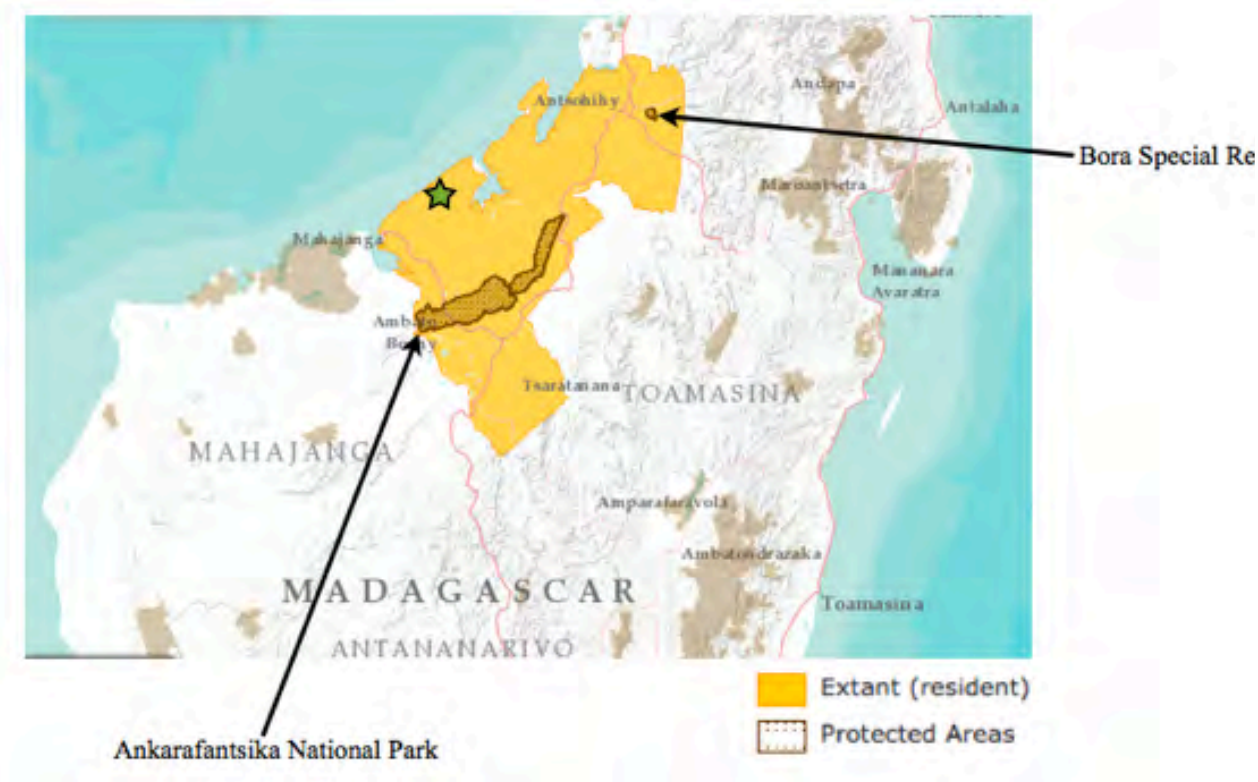


Figure 1.1: *P. coquereli* is a species exhibiting micro endemism, found solely in the North West of the island of Madagascar, depicted by the area in yellow. The need to increase protected areas within its habitat is great, as seen by the insignificant stretches of land protected at present time. These are Ankarafantsika National Park and Bora Special Reserve. The Mariarano village, where my research was carried out, is represented by the green star and is found in the North West of the island, with dry deciduous forest.[13](Source: IUCN Red List, adapted by A. Baranowski, July 2014)

[13] http://maps.iucnredlist.org/map.html?id=18355

HYPOTHESES

2.1 TREE HEIGHT

From my time spent observing *P. coquereli* in the preliminary opportunistic surveys and research carried out prior to visiting Madagascar, I believe they will be selective in their tree use, preferring taller trees (15-20m) reaching above the forest canopy to facilitate free locomotion by "vertical clinging and leaping"[1]. Avoiding ground predators found in the area, such as the fossa, would also be a reason. Threatened by large predatory birds such as hawks, I expect the sifakas to avoid overly exposed areas, and so the tallest trees (>25m).

2.2 TREE SPECIES

Tree species serving *P. coquereli* a purpose, such as bearing fruit, will be preferred. Since research took place in the dry season in the absence of fruit, further research could determine the presence of a correlation between the phenology of tree species and selectivity. The two factors I chose to investigate could be linked if certain species reached certain heights, for example. This could lead to selectivity being affected by multiple factors simultaneously, providing unreliable results. Tree height is expected to be more significant in determining the selectiveness of trees by *P. coquereli*.

[1] Only leg power is used, and an upright position is maintained whilst jumping from tree to tree.

METHODOLOGY

3.1 OPPORTUNISTIC SURVEYS – PRELIMINARY WORK

Choosing a suitable method was intrinsic to obtain precise results in a time-efficient manner and to ensure reliability. Plot-based methods[1] generally tend to be more time-consuming and prone to observer bias. *P. coquereli* are found in motion and above ground level, accentuating these disadvantages and so leading me to choose a plot-less method called Point Centered Quarter (PCQ) *(see **section 3.4**).* Several opportunistic surveys were necessary before commencing recorded surveys, to see the suitability of the method for my investigation. These consisted of walking along each of the five routes at Mariarano and spotting *P. coquereli,* performing the PCQ method to get acquainted with it, and recording the preliminary results. Though initially data collection was to occur in two separate camps, it was seen that the forest type in the second resembled that of wetlands, and so was different enough from Mariarano for inaccuracies to arise in the results. By then deciding to constrict the scope of the investigation to the dry deciduous forest in Mariarano, the effect of tree height and species could be seen on the selectivity of *P. coquereli* in their tree use, rather than that of forest type.

3.2 TRANSECT WALKS

Surveys were performed in the mornings (starting at 7:30am), afternoons (starting at 1:30pm) and evenings (starting 7:30pm) by line transect sampling, whereby observers walked along each of the transects at a slow and steady pace.[2] A local guide accompanied observers, leading the way and providing the vernacular names of the trees in Malagasy. The width of each transect was

[1] Quadrat sampling is an example, where a specific area is observed for a set period of time.
[2] Approximately $3kmh^{-1}$.

between two and three metres, and lengths ranged from 2.5km to 5km. When a group of *P. coquereli* was spotted, the perpendicular distance of the tree from the path was recorded and the PCQ method was carried out *(see **section 3.4**).* The transect walk would then continue.

All routes had several distinctive characteristics; dense and dry vegetation, arid landscape, thin-barked trees with low canopies, little sunlight, abundance of dead wood and leaves on the forest floor. Disturbance was intermediate in all routes. Due to the physical similarities, their proximity and the results obtained from preliminary surveys, it was assumed that the routes were similar enough for data obtained from each to be combined into a collective dataset. Transect walks consisted of three individuals, two of which surveyed for *P. coquereli,* with a third recording observations and data from PCQs. Each route was walked at least six times throughout the investigation to ensure reliability of the results.

3.3 POINT OF COMPARISON – RANDOM TREES

To compare the results of the trees selected by *P. coquereli* and see whether tree height and tree species affected this, eighty random trees were selected from the five routes of the Mariarano area, with sixteen trees from each route. During opportunistic surveys, random trees were selected by generating random co-ordinates on a GPS device. The nearest tree (DBH ≥ 5cm) from those co-ordinates was the central tree; PCQ was carried out on all random trees.

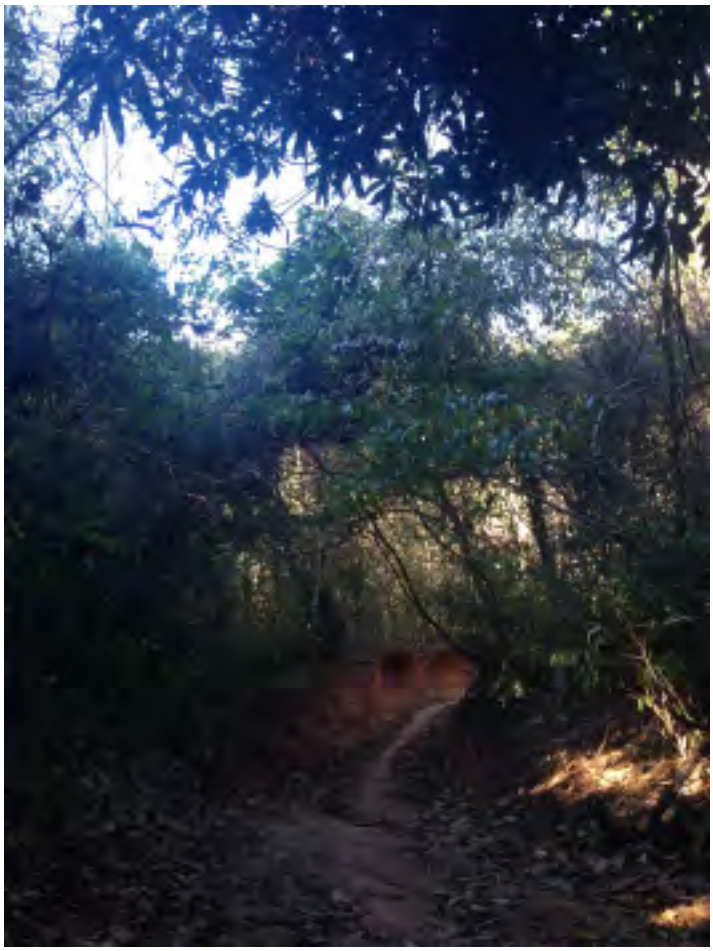


Figure 3.1: Image of Route 4 at Mariarano, depicting the arid, dense vegetation, low canopies and thin branches trees described in **section 3.2 (**Baranowski, July 2014)

3.4 PCQ

The Point Centred Quarter method[3] (PCQ) is a plot-less sampling method, so is not carried out in a specific known area over a period of time; in our case, trees selected by *P. coquereli* were recorded along the Mariarano transects. PCQ was used as it provides information on four surrounding trees in addition to the tree selected by *P. coquereli,* providing a more holistic image of their habitat use

[3] Mitchell (2007)

whilst avoiding inefficiencies of plot-based methods *(see **section 3.1**).*

The sampling method was considerably straightforward, requiring a maximum of three individuals, as opposed to more time-consuming plot-based (quadrat) methods. Observation took place by transect line sampling *(see **section 3.2**),* whereby individuals walked along each transect spotting *P. coquereli.*When a group of *P. coquereli* was spotted on a tree, the perpendicular distance of that tree[4] from the path and group size of *P. coquereli* were recorded, as were the GPS co-ordinates. For each central tree, the diameter at breast height (DBH), height, canopy height, species and phenology were recorded.

Subsequently, the nearest trees in the NE, NW, SE and SW directions with DBH ≥ 5cm were selected, and their distance from the central tree was recorded *(see **Fig 3.2**, next page).* Measurements of DBH, height, canopy height, species and phenology were also obtained for the four quarter trees.

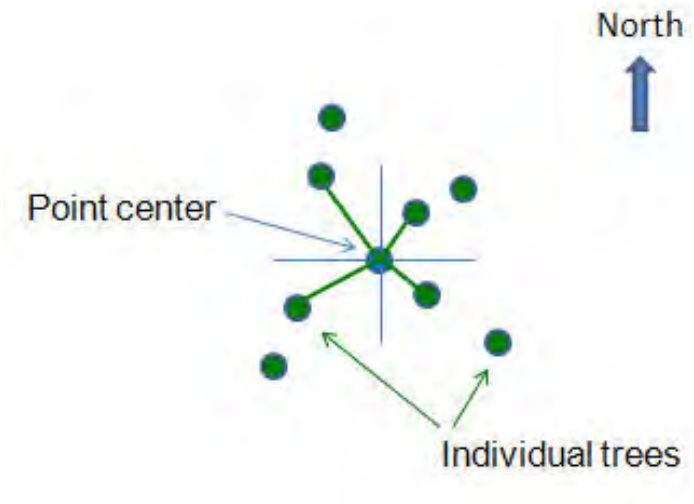


Figure 3.2: Illustration of PCQ method depicting the central tree and four nearest trees in the NE, NW, SE and SW directions, as well as other trees found further away. All trees are depicted by green dots.[5]

[4] The tree on which the sifaka sighting occurred is referred to as the central tree.
[5] http://faculty.wwu.edu/wallin/envr407/407_PCQ_data_and_lidar_lab.html

RESULTS

4.1 DESCRIPTIVES AND STATISTICAL ANALYIS OF DATA CONCERNING THE EFFECT OF TREE HEIGHT

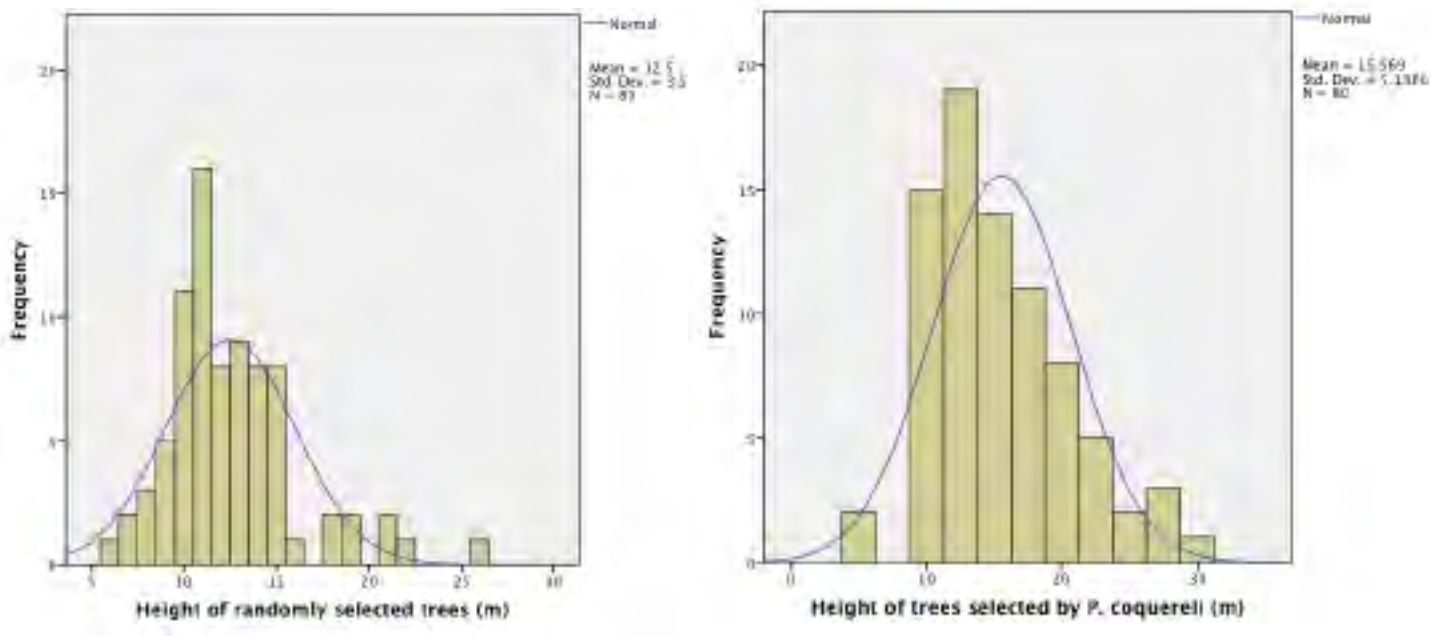
	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness
	Statistic							
Height of randomly selected trees (m)	80	20	6	26	12.500	3.515	12.354	1.286
Height of trees selected by *P. coquereli* (m)	80	24	5	29	15.569	5.1366	26.384	.705

	Skewness	Kurtosis	
	Std. Error	Statistic	Std. Error
Height of randomly selected trees (m)	.269	2.592	.532
Height of trees selected by *P. coquereli* (m)	.269	.146	.532

Table 4.1: In the table above, N represents the sample size for randomly selected trees and those selected by *P. coquereli.* The standard deviation and variance are measures of spread, with 68% of all data values falling ±1 standard deviation from the mean and 95% found ±2 standard deviations. The variance is the square of the standard deviation. Skewness and kurtosis compare the spread of data to the Normal Distribution.

Skewness measures spread of data to the left or right of the mean, and a value larger than ±1 is said to be significantly different from a normal distribution. Similarly, kurtosis measures if data is clustered around the mean or spread sparsely, with a value larger than ±1 signifying the lack of a Normal Distribution. SPSS software and Excel were used to perform statistical tests and generate tables and graphs.

Graphs showing the distribution of the two groups of eighty trees in comparison with the Normal Distribution



Figures 4.1 & 4.2: Depiction of the distribution of randomly selected trees and trees selected by *P. coquereli.* Here it can be seen clearly that the graph of randomly selected trees is slightly skewed to the right, as described in **Table 4.1** (g_1=1.286)[1]. This graph is also leptokurtic[2] (g_2= 2.592)[3], signifying that a large proportion of the data is clustered around the mean height. The graph of the trees selected by *P. coquereli* follow the curve of the Normal Distribution within the acceptable limits aforementioned *(see previous page)* (g_1=0.705, g_2 = 0.146 *see* **Table 4.1**).

[1] g_1 being the estimated skewness parameter
[2] Clustered around the mean
[3] g_2 being the estimated kurtosis parameter

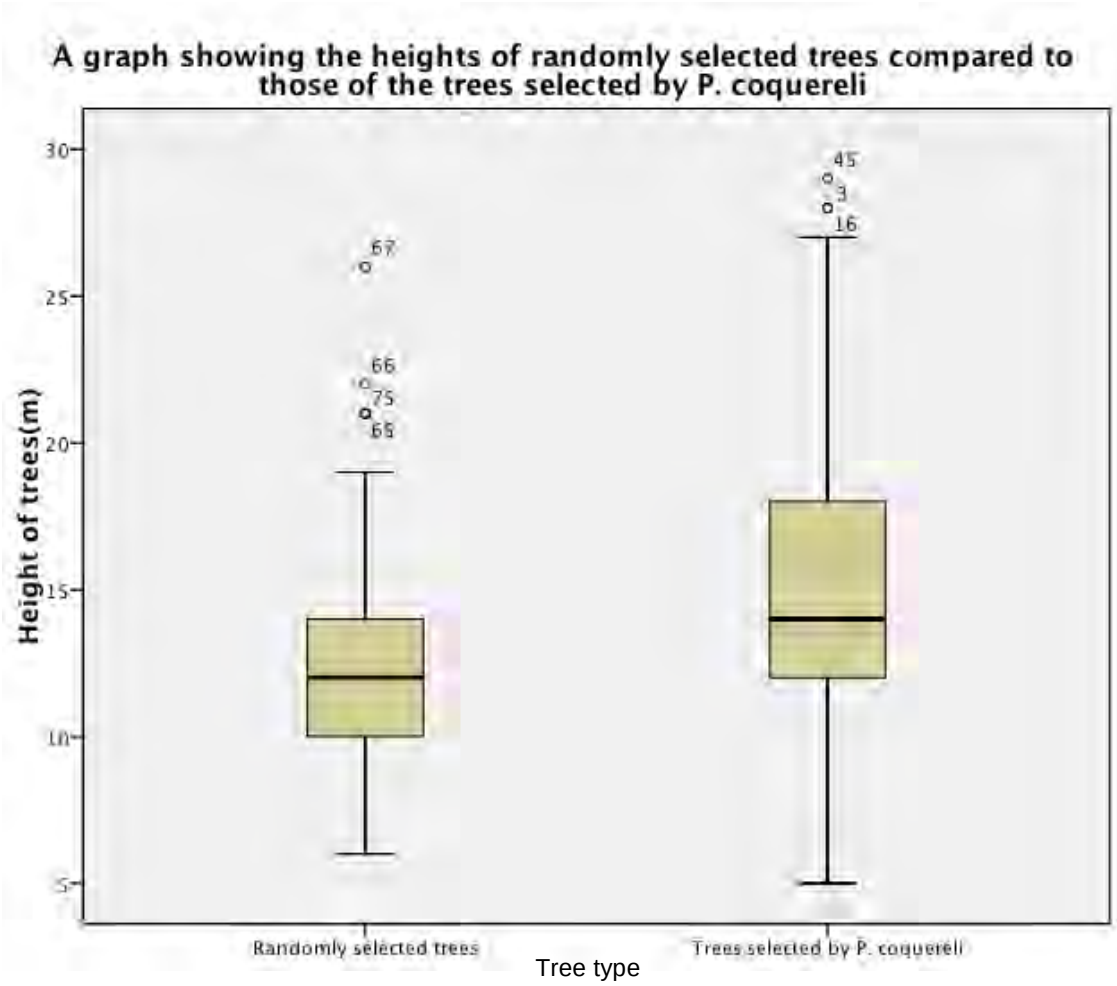


Figure 4.3: Box plots illustrating the height of trees for the eighty randomly selected trees compared to the eighty trees selected by *P. coquereli.*

The whiskers in **Fig 4.3** depict the range of heights with the outliers[4] being shown as circles, and the median as the horizontal line within the boxes. These represent the interquartile range, and their overlap could suggest no significant difference between the heights. However, the trees selected by the sifakas reach much greater heights within their range, and a larger proportion of trees used by them are taller. This was further investigated using just the mean heights *(see **Fig 4.4**).*

[4] Values found at the extreme ends of the range, outside 3 standard deviations (which have a 99.7% CI).

A graph showing the mean heights of randomly selected trees and trees selected by *P. coquereli* with error bars covering a 95% Confidence Interval (CI)

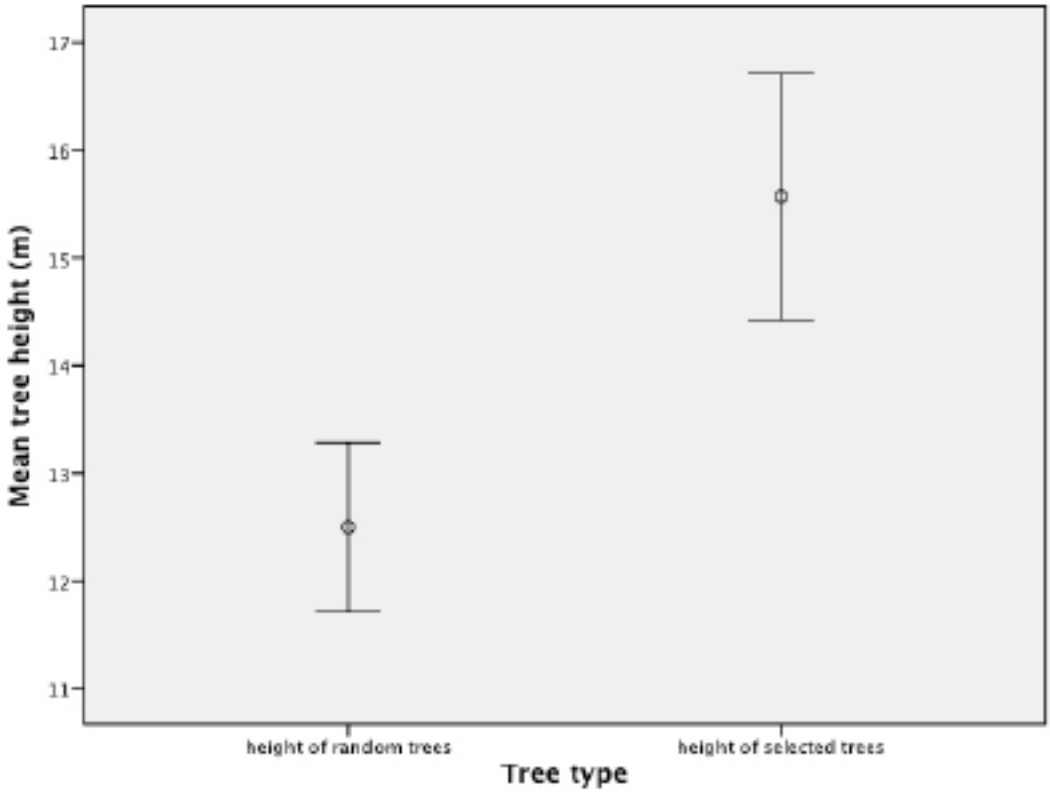


Figure 4.4: Graph illustrating the mean heights of the two groups, depicted by the circles; the whiskers are error bars extending to cover a 95% confidence interval (CI)[5]. The error bars do not overlap, suggesting that there is a significant difference between the average height of randomly selected trees and those selected by *P. coquereli.*

Performing an independent samples t-test (*see app. B.1),* a significant difference was found between the mean heights of selected and random trees ($|t_{158}|= 4.41$, $t_c=1.99$, $p<0.05$). The t_c (critical t-value) is less than the calculated t-value $|t_{158}|$, showing a significant difference. The number 158 represents the degrees of freedom, calculated by subtracting two from the total sample size. This led to the rejection of the null hypothesis, stating that there was no significant difference. As

[5] The mean height is expected to be found within these boundaries with 95% confidence.

the heights of random trees did not seem to be distributed normally, I performed a Levene's test for equality of variances, which returned a significantly high difference in the variances of the two tree groups (F=12.75, sig.=0.00)[6]. As one of the conditions for a successful t-test is equal variances, this led to doubting the outcome of the first test. Therefore, a non-parametric[7] test in the form of the Mann Whitney U was performed (*see app. B.2)* to ensure the reliability of the t-test results. Again, a significant difference between the two samples was found (U=1977, z=-4.193, p<0.05)[8].

4.2 RESULTS CONCERNING THE EFFECT OF TREE SPECIES

Trees selected by *P. coquereli*	
Frequency	Species
6	*Securinega seyrigii*
6	*Trilepisium sp.*
5	*Albizia sp.*
4	*Chrysophyllum sp. 1*
4	*Dalbergia mollis*
4	*Hyperacanthus perrieri*
4	*Milletia richardiana*
4	*Rothmania sp.*

Table 4.2

Randomly selected trees	
Frequency	Species
9	*Dalbergia mollis*
7	*Grewia grandidieri*
7	*Rothmania sp.*
7	*Trilepisium sp.*
5	*Commiphora coleopsis*
3	*Milletia richardiana*
3	*Securinega seyrigii*
3	*Tina isaloensis*

Table 4.3

Tables 4.2 & 4.3: The two tables depict the eight species of tree with the highest frequencies, for trees selected by *P. coquereli* and randomly selected trees. As seen from **table 4.2,** the trees selected by *P. coquereli* have a bimodal distribution (there are two modal species) whilst randomly selected trees are distributed in a unimodal manner, and so one species of tree is most abundant. Though there is no striking correlation between the frequencies of randomly selected trees and

[6] If sig<0.05, then the variances are not deemed to be equal.
[7] Not taking into account the distribution of data or the variance.
[8] If p<0.05, H_0 =there is no significant difference, is rejected.

those chosen by *P. coquereli*, it can be seen that the abundance of tree species in the forest plays a role to which trees are selected by the sifakas from the fact that numerous species present in **table 4.3** can also be seen in **table 4.2**. This weak correlation could suggest that tree species is not an important factor affecting the selectiveness of trees by *P. coquereli.*

Table showing phenology, uses and stating endemism of eight plant species most commonly used by *P. coquereli*

Species	Phenology	Human Uses of Plant	Endemic to Madagascar
Securinega seyrigii[9]	Flower buds in early spring, no fruit, leaves.	Medicinal plant, construction material	Yes (South-Western Madagascar)
Trilepisium sp.	Leaves, fruit	No information.	Yes
Albizia sp.[10]	Bears fruit in early-mid spring, leaves.	Exports of timber	No
Chrysophyllum sp. 1	Leaves, no further information.	Construction material	Yes
Dalbergia mollis[11]	Flowers August-October, bears fruit.	Type of rosewood, used for exports, logged.	Yes
Hyperacanthus perrieri[12]	Leaves, flowers.	No information	Yes
Milletia richardiana[13]	Flowers August-December, bears fruit.	Furniture, construction materials	Yes (Western Madagascar)
Rothmania sp.[14]	Leaves, no further information.	Medicinal plant	Yes

Table 4.4: Table showing the phenology and uses of the eight most commonly used trees by *P. coquereli,* also stating whether each species is endemic to Madagascar (and if it exhibits micro-endemism). Most species bear fruit, and are endemic, but many are also logged in order to be used as construction materials, for furniture, musical instruments and to be exported overseas, posing a great threat by destroying the habitats of endangered species such as *P. coquereli.*

[9]Charles-Dominique (1980), Boiteau *et al.* (1999)
[10]Charles-Dominique (1980)
[11]Du Puy *et. al* (2002)
[12] http://data.gbif.org/species/5596946/
[13]Du Puy & Labat
[14] Iwu

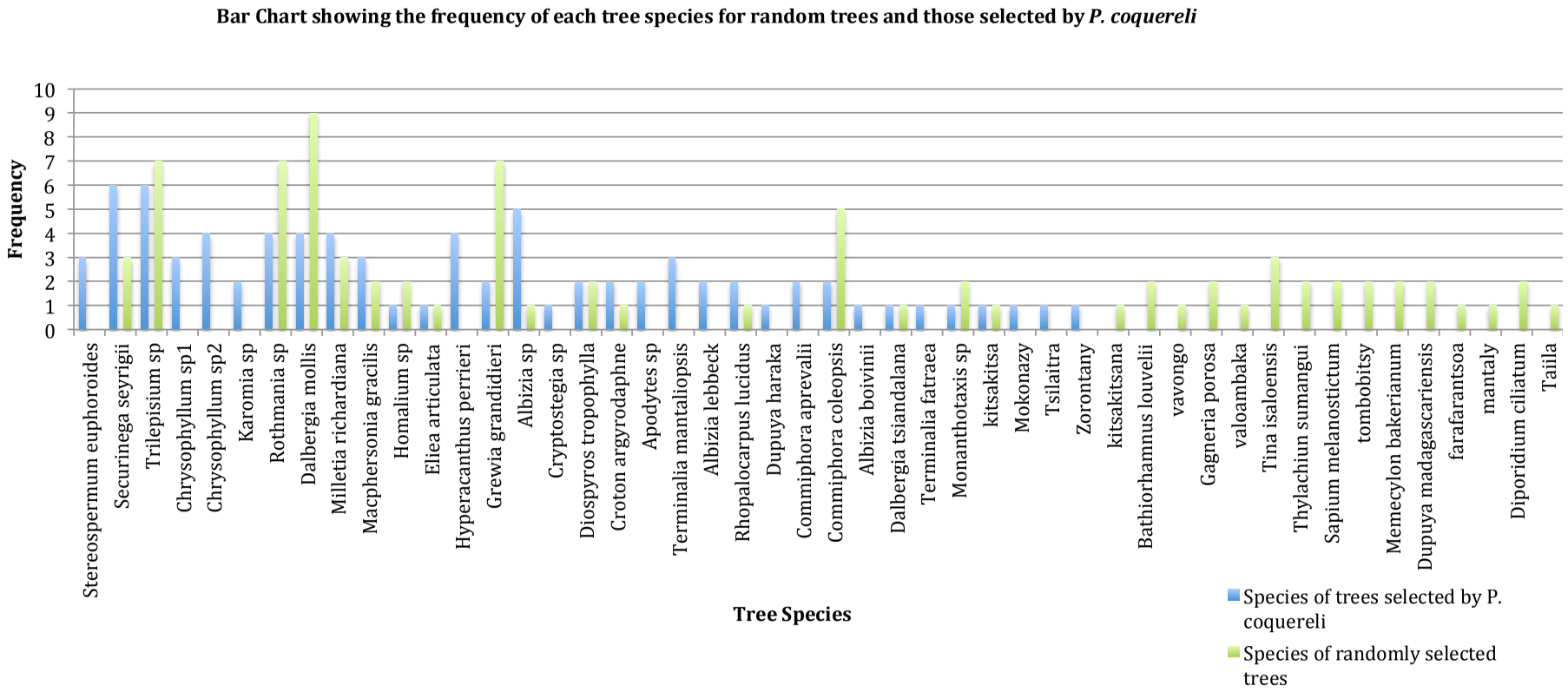


Figure 4.5: Bar chart showing the frequency of each species of tree present in each of the two groups. The blue bars depict the trees selected by P. coquereli, and the green bars the randomly selected trees. Of the eighty trees sampled for each group, 33 different species were present in trees selected by *P. coquereli,* and 40 in the randomly selected trees. It can be seen that there is not a noteworthy correlation between the abundance of a species and its use by *P. coquereli.*

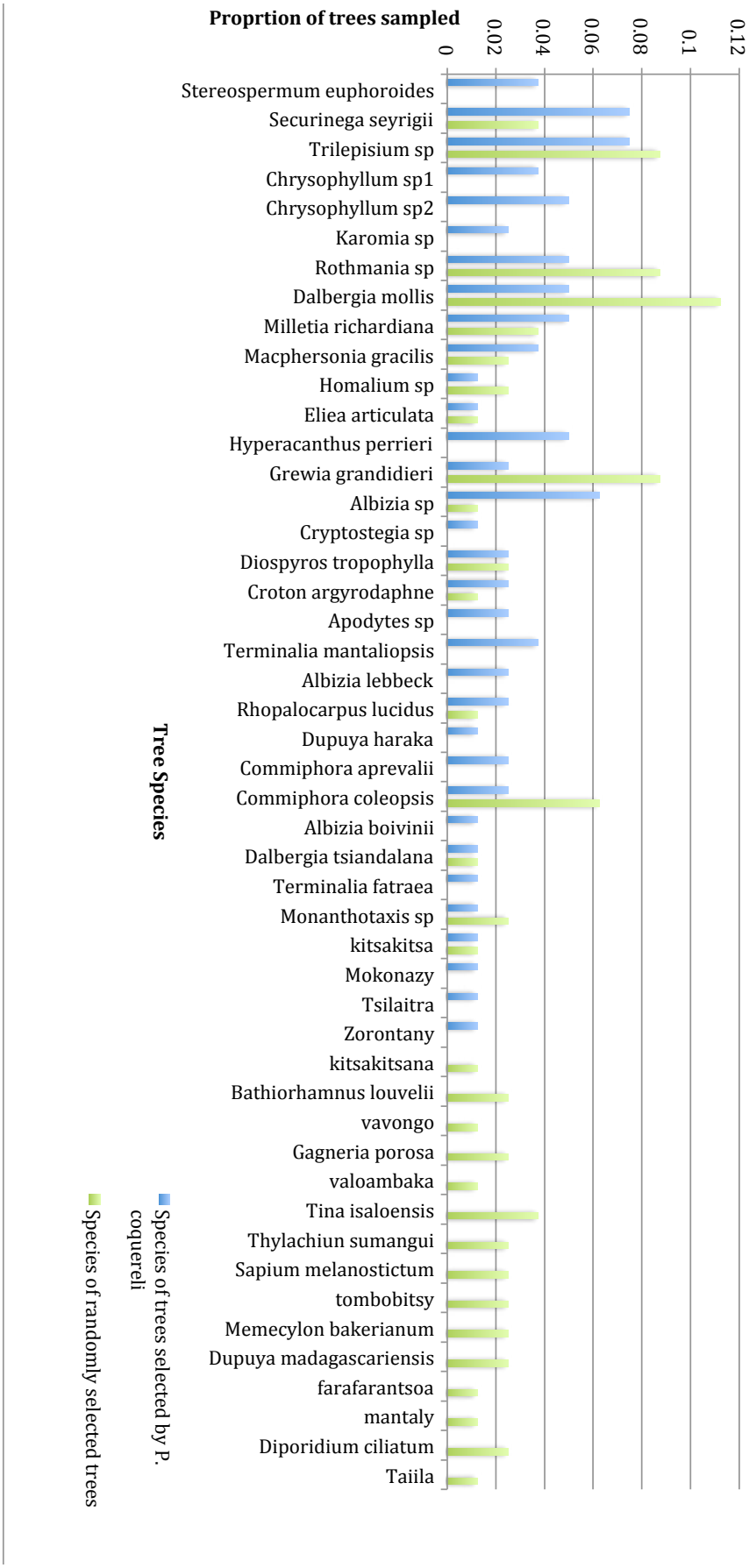


Figure 4.6: This bar chart illustrates each species of tree as a proportion of all trees sampled for each group. Again, the blue bars depict the trees selected by P. coquereli and the green the randomly selected trees. The trees selected by *P. coquereli* are distributed bimodally, with the species *Trilepisium sp.* and *Securinega seyrigii* each comprising 7.5% of the total sample. The randomly selected trees show a unimodal distribution, with the species *Dalbergia mollis* being the most abundant.

DISCUSSION

The research questions read “To what extent are Coquerel’s sifakas (*Propithecus coquereli)* selective in their tree use based on tree height in the dry deciduous forest of Mariarano, Northwest Madagascar? And further, to what extent are *P. coquereli* selective in their tree use based on species of tree?”

5.1 TREE HEIGHT

The mean height of selected trees was 15.6m whilst that of random trees was just 12.5m, supporting the hypothesis.[1] From **Figs 4.1 & 4.2** we see that though the range of heights is similar for the two groups *(see **table 4.1**),* their distribution presents several dissimilarities. Figure **4.3,** however, illustrates that the similarity in height range was misleading, as outliers present in the random trees led to the range being depicted as greater than it was. Due to the right skewness of the graph[2] of randomly selected trees ***(Fig 4.1)*** the mean height is towards the tail, showing most trees to be below the average height of 12.5m *(see **Fig 4.1**).* There is also a discontinuity in the height distribution on the descending side of the curve, with height frequency decreasing sharply above 15m. The heights of the trees selected by *P. coquereli,* contrastingly, display a considerably normal distribution (g_1=0.705, g_2 = 0.146, *see* **Table 4.1**). Very few trees were shorter than 10m (see **Fig 4.2**), with most being found above the modal height of 13m. This difference in distribution suggests selective behaviour, by sifakas preferring taller trees. The average canopy height of central trees selected by sifakas was 8.9m whilst that of the random trees was just 3.3m, further supporting the initial hypothesis; it suggests that by selecting trees extending above the forest canopy, sifakas will have space for locomotion, feeding and social interaction.

[1] This stated that on average *P. coquereli* would use the forest’s taller trees.
[2] g_1=1.286, (If g_1>1 then curve deemed to be skewed, see section 4.1)

The maximum height of trees selected by *P. coquereli* was 29m, and of random trees 26m. Though both groups have low frequencies of trees >25m, **Figs 4.1 & 4.2** show an unexpected result. In the trees selected by *P. coquereli,* the frequency of trees <10m is much lower than that of trees >25m, whilst in the group of randomly selected trees there is a progressively increasing number of trees measuring <10m and very few instances of trees >25m. This incongruence between the two graphs does not fully support the hypothesis,[3] suggesting that ground predators such as fossas or boa constrictors pose a larger threat to *P. coquereli* than airborne predators. This leads to the avoidance of shorter trees, even risking exposure to raptors such as hawks Wright's work supports this, deeming that lemurs exhibit selective behaviour to avoid predation by animals such as the fossa.[4] However, the lack of space or suitable branches may have also caused the very few sifaka sightings on shorter trees, as may have decreased visibility of the observers due to denser vegetation.

In Fig 4.4 the whiskers of the two groups (which cover a 95% CI) do not overlap, insinuating a significant difference between randomly selected trees and those preferred by *P. coquereli*. To test this, I performed an independent samples t-test (*see app. B.1),* finding a significant difference between the mean heights of selected and random trees ($|t_{158}|= 4.41$, $t_c=1.99$, $p<0.05$). Due to a considerable difference in the variances of the two groups[5] and the slight skewness of the group of random trees[6] I decided to ensure the reliability of the t-test results through a non-parametric test[7]. Using a Mann Whitney U (*see. B.2 for explanation)* I once again found a significant difference ($U=1977$, $z=-4.193$, $p<0.05$)[8]. Though usually not as powerful as a t-test, an advantage was that the

[3] This predicted the lowest frequency of the selected trees to be the tallest.
[4] Wright (1998)
[5] Levene's test: F=12.75, sig.=0.00 (see **table 4.1**)
[6] $g_1= 1.286$ (see **table 4.1**)
[7] A test which would not take into account the differences in variance or lack of normal distribution of the random tree group.
[8] $p<0.05$ so H_0 rejected and difference deemed significant.

outliers from the random tree group *(see **Fig 4.3**)* would not distort the results, as all values are assigned a rank prior to performing the test from the lowest to highest values. The results of the statistical tests support the hypothesis to a great extent, showing that *P. coquereli* are selective in their tree use based on tree height.

The implications of these findings are exceedingly important, as they enrich the databases providing data on *P. coquereli.* By showing that *P. coquereli* prefer trees that are on average 3m taller than randomly selected trees, a foundation for commencing the replantation of lost forest has been set, aiding conservationists to recreate habitats enabling this species to survive and reproduce. However, tree height is not the sole factor affecting the selection of trees by Coquerel's sifakas, and thus the effect of tree species was also investigated.

5.2 TREE SPECIES

The randomly selected trees consisted of 40 different species of tree, whilst trees selected by *P. coquereli* were of 33 different species. This is not greatly significant as numerous species of trees were selected only once by the sifakas (*see **Fig 4.5***), suggesting random behaviour. The hypothesis stated that sifakas would be selective if the particular species served them a purpose, for example by bearing fruit or having better tasting leaves. Further investigating the phenology of the eight most frequently used trees *(see **table 4.4**)* showed that most trees selected by *P. coquereli* bear fruit, supporting the hypothesis. Due to fieldwork being carried out in the dry season and so the absence of fruit, the particular plants would not have been directly useful; they can, however, perhaps be recognised by sifakas by other features such as their leaves and still selected over non fruit-bearing trees.

It was interesting to see that trees were not selected by the sifakas proportionally to their abundance *(see **Fig 4.5** and **tables 4.2 & 4.3**).* For example, *Dalbergia mollis* was most abundant comprising 9 of the 80 random trees, but was only fourth most popular amongst *P. coquereli.* Contrastingly, many species selected by *P. coquereli* are either not abundant or not even present in the randomly selected trees. These results agree with the hypothesis, having predicted the effect of tree species to be less powerful than that of tree height. Another significant finding was to see that a large proportion of the trees most used by *P. coquereli* is logged or exported. Though exports are the largest source of income of the country and often vital for the livelihood of locals, the threat posed to the habitats of these endangered animals is immense. Conservationists must resolve this issue for sustainable development to proceed in Madagascar without sacrificing biodiversity.

5.3 LIMITATIONS

5.3.1 LIMITATIONS OF METHODOLOGY

Research was carried out in the Malagasy dry season, so no fruit were present on trees, hindering the ability to determine reasons behind tree use of specific species. Replication of the investigation could occur over the period of a year to ensure that the results are not affected by seasonal changes.

During transect walks, it was assumed that all sifakas found 0m from the transect line were spotted, which cannot be certain. The effect of human presence on that of *P. coquereli* must also be considered. However, this particular species does not seem to be greatly affected by human presence, often being spotted in the trees above the base camp. It was further assumed that all routes consisted of the same forest type and so would provide the same habitat for *P. coquereli.* To

ensure this, surveys could be carried out on the vegetation, and the degradation determined in each route.

The PCQ method has been shown to be more efficient in producing accurate estimates of species density whilst also providing more data in comparison with plot-based and plot-less techniques, such as the nearest neighbour or closest individual methods.[9] However, several limitations must be considered. The height of trees was often an estimate, it being exceedingly inefficient and time-consuming to measure each tree's height. As observations occurred in groups of three, the same individual did not carry out the process each time, possibly leading to Individual bias and incongruence. By having a larger dataset, this effect could be minimised.

5.3.2 LIMITATIONS OF RESULTS

The large difference in variance between the heights of the two groups and skewness of the curve of random trees hindered the effect of the t-test. This was overcome by performing the non-parametric Mann-Whitney U, which did not take data distribution into consideration, and minimised the effects of outliers by converting results to ranks prior to the test.

The dataset was considerably small and should be expanded when repeating the study, particularly considering that only sixteen trees from each route were present in each group. The outliers in **Fig 4.3** could lead to misjudgment of data and the conclusion that there is no significant difference between the tree heights of the two groups. However, **Fig 4.4** depicts the significant difference with clarity.

In certain cases the scientific names of particular species could not be found, so their vernacular names were used. This led to a lack of uniformity within the

[9] Beasom & Haucke (1975)

results, as did the absence of information on the phenology or uses of certain species. Increasing the scale of the investigation so that these cases form a smaller proportion of the dataset could minimise this effect.

CONCLUSION

This investigation sought to determine whether Coquerel's sifakas (*P. coquereli*) are selective in their tree use through analysing two affecting factors – tree height and tree species, to answer the research questions reading "To what extent are Coquerel's sifakas (*Propithecus coquereli)* selective in their tree use based on tree height in the dry deciduous forest of Mariarano, Northwest Madagascar? And further, to what extent are *P. coquereli* selective in their tree use based on species of tree?" Through analysing the data collected it was found that there was a significant difference between the heights of trees selected by *P. coquereli* and randomly selected trees, with those selected by the sifakas being, on average, 3m taller; this evidence supported the initial hypothesis, as well as enriching the data available on habitat preferences of *P. coquereli,* a species which has not received thorough research. Concerning the effect of tree species on the selection of trees, of the most used species the majority bore fruit, all but one were endemic to either Madagascar or the region, and many are threatened by logging. These results were less significant than those concerning tree height, possibly due to research being undertaken in the dry season. Unexpectedly, many of the species of trees selected by *P. coquereli* were not present in the random selection, perhaps suggesting a preference.

FURTHER INVESTIGATION

Though answering the research questions to a great extent and obtaining useful results, there are still unresolved issues needing to be investigated to develop the most effective conservation strategies possible.

The effect of ground predators such as the fossa as a direct threat to *P. coquereli* could be investigated, and by combining the results the most important factors

leading *P. coquereli* to be selective in their tree use could be seen. Furthermore, to ameliorate results concerning species of tree, a longitudinal experiment could be conducted over the course of a year. By taking into account the differences arising with the variation of the seasons, a more profound conclusion can be reached as to whether the species of tree affects selectivity of trees by *P. coquereli.*

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Ms. Hélène Bonsall, for her continued supported throughout my planning, preliminary work, fieldwork and execution of my essay, providing her expertise and utmost support. I am also exceedingly grateful to Mrs. Kerry Pitcher, Head of Biology, for organising the trip to Madagascar and making this opportunity possible. I would like to express my gratitude to Barry Ferguson of Operation Wallacea in Madagascar for providing me with an insight on the Malagasy culture and way of life without which I would not have truly understood importance of conservation work, and for being more than happy to answer any queries with great patience throughout the execution of my essay. I would finally like to thank Remi for helping me carry out my fieldwork, and the local guides for their expertise that proved to be invaluable.

BIBLIOGRAPHY

Arkive.org, (2014). *Coquerel's sifaka photo - Propithecus coquereli - G1568 | ARKive*. [online] Available at: http://www.arkive.org/coquerels-sifaka/propithecus-coquereli/image-G1568.html [Accessed 19 Jul. 2014].

Beasom, S. and Haucke, H. (1975). A comparison of four distance sampling techniques in South Texas live oak mottes. *Journal of Range Management*, pp.142--144.

Black, R. (2012). Lemurs Sliding Toward Extinction. *BBC*. [online] Available at: http://www.bbc.co.uk/news/science-environment-18825901 [Accessed 16 Jul. 2014].

Du Puy, D. (2002). The Leguminosae of Madagascar. *Curtis's Botanical Magazine*, 14(4), pp.231--240.

Explore and discover Red List species ranges and observations, (2014). *IUCN Red List maps*. [online] Available at: http://maps.iucnredlist.org/map.html?id=18355 [Accessed 12 Jun. 2014].

Ferguson, B. (2014). *Biodiversity in Madagascar*.

Ferguson, B. (2014). *Conservation Methods in Madagascar*.

Ferguson, B. (2014). *Culture and Livelihoods in Madagascar*.

Grieser, B. (2014). Infant Development and Paternal Care in Two Species of Sifakas. *Primates*, 33(3), pp.305-314.

Hladik, C., Charles-Dominique, P., Petter, J. and others, (1980). Feeding strategies of five nocturnal prosimians in the dry forest of the west coast of Madagascar. *Nocturnal Malagasy primates: ecology, physiology, and behavior*, pp.41-73.

Isalo Rock Lodge Madagascar. (2013). *Some Hints For Successful Eco Tourism in*

Madagascar. Available: http://www.isalorocklodge.com/eco-tourism.html [Accessed 20 Jun. 2014].

Iucnredlist.org, (2014). *Propithecus coquereli (Coquerel's Sifaka, Crowned Sifaka)*. [online] Available at: http://iucnredlist.org/details/18355/0 [Accessed 26 Jul. 2014].

Iwu, M. (2013). *Handbook of African Medicinal Plants, Second Edition*. 1st ed. Hoboken: Taylor and Francis

Lemur.duke.edu, (2014). *Coquerel's Sifaka | Duke Lemur Center*. [online] Available at: http://lemur.duke.edu/discover/meet-the-lemurs/coquerels-sifaka/ [Accessed 3 Jun. 2014].

Mitchell, K. (2010). Quantitative analysis by the point-centered quarter method. *arXiv preprint arXiv:1010.3303*.

Mittermeier, R. and Nash, S. (2006). *Lemurs of Madagascar*. 1st ed. Washington, DC: Conservation International.

Mittermeier, R., Schwitzer, C., Davies, N. and Johnson, S. (2013). *Lemurs of Madagascar: A Strategy For Their Conservation 2013-2016*. 1st ed. Bristol: IUCN SSC Primate Specialist Group, Bristol Conservation and Science Foundation, and Conservation International, pp.5-11, 35-47.

Myers, N. (1988). Threatened biotas:" hot spots" in tropical forests. *Environmentalist*, 8(3), pp.187-208.

Petter, J. (1962). Ecological and behavioral studies of Madagascar lemurs in the field. *Annals of the New York Academy of Sciences*, 102(2), pp.267--281.

Worldbank.org, (2014). *Madagascar: Measuring the Impact of the Political Crisis*. [online] Available at: http://www.worldbank.org/en/news/feature/2013/06/05/madagascar-measuring-the-impact-of-the-political-crisis [Accessed 24 Jul. 2014].

Wright, P. (1999). Lemur traits and Madagascar ecology: coping with an island environment. *American journal of physical anthropology*, 110(S29), pp.31--72.

WWF, (2014). *Pioneering project against deforestation*. [online] Available at: http://wwf.panda.org/what_we_do/where_we_work/project/projects_in_depth/conservation_program2/ [Accessed 11 Jun. 2014].

APPENDIX A

COQUEREL'S SIFAKA (*PROPITHECUS COQUERELI)*: INFORMATION

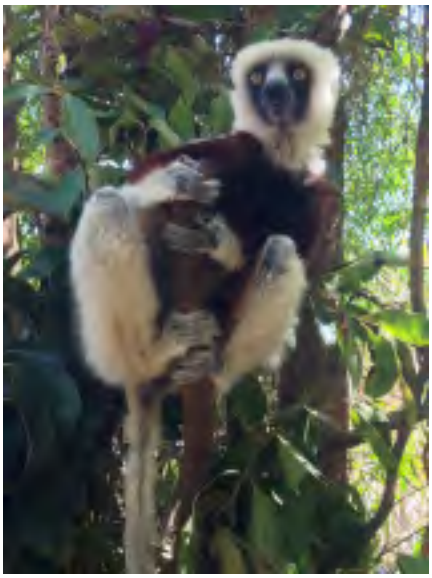


Figure : *P. coquereli* (Baranowski, July 2014)

Family: Indriidae

Weight: 3.7-4.5kg[1]

Height: 92.5-110cm[2]

Diet: Primarily young leaves, bark, fruit, dead wood in the wet season and dry leaves in the dry season.[3] Foraging will normally comprise 30-40% of their daily activities.[4]

Characteristic features: The name 'sifaka' comes from call they make, which is *shi-fak* in Malagasy.[5] The sifakas move from tree to tree only using the strength of back legs, in an upright position described as 'vertical clinging and leaping' and can travel up to just over 6m in the air from tree to tree. Young able to give birth at 3.5 years, one baby per year on average. Can be distinguished by

[1] http://eol.org/pages/2925752/details
[2] http://www.arkive.org/coquerels-sifaka/propithecus-coquereli/image-G1568.html
[3] Ferguson (2014)
[4] http://lemur.duke.edu/discover/meet-the-lemurs/coquerels-sifaka/
[5] Ferguson (2014)

different shade of brown-dark. Males' chests, arms and thighs are predominantly darker.

Threats: A fady (taboo) on eating & killing lemurs had protected the endangered species till now but recent immigrants do not respect these. Immigration also leads to hunting, farmers set fires to forests to create more land for their animals, logging for charcoal and firewood and slash-and-burn are also threats.

Social Behaviour: They live in groups of 3-10 although most often in about 4-5,[6] females dominant to males, home range of 4-9 hectares.[7]

Species Range: They are endemic to Madagascar and are found in the dry decidous forests of the North West of the island, to the North and East of the Betsiboka river.[8]

IUCN rating: Endangered (since 1996)

[6] Ferguson (2014)
[7] Petter (1962)
[8] Mittermeier *et al.* (2006)

APPENDIX B

STATISTICAL ANALYSIS

B.1 T-TEST

Group Statistics

	Tree Type	N	Mean	Std. Deviation	Std. Error Mean
Height of trees	Selected trees	80	12.50	3.515	.393
	Random trees	80	15.57	5.137	.574

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig.	t	df	Sig. (2-tailed)
Height of trees	Equal variances assumed	12.745	.000	-4.410	158	.000
	Equal variances not assumed			-4.410	139.679	.000

		t-test for Equality of Means			
		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
				Lower	Upper
Height of trees	Equal variances assumed	-3.069	.696	-4.443	-1.694
	Equal variances not assumed	-3.069	.696	-4.445	-1.693

t_c (158 d.f., p=0.05) = 1.9751

t_c (140 d.f., p=0.05) = 1.9775

$t > t_c$ so H_0 is rejected.

According to Levene's test for the equality of variances, if the significance is <0.05 then there is no equality in the variances. Therefore, the bottom row of figures of the t-test must be used, which take this into consideration. In this case, the calculated t-value is -4.410, which is greater than the critical t value for 140 degrees of freedom (less degrees of freedom are used to take the inequality of the variances into account). Therefore, there is a significant difference between the heights of the randomly selected trees and those selected by *P. coquereli.*

B.2 MANN - WHITNEY U TEST

Ranks

	Tree type	N	Mean Rank	Sum of Ranks
Height of trees	Random trees	80	65.21	5216.50
	Selected trees	80	95.79	7663.50
	Total	160		

Test Statistics[a]

	Height of trees
Mann-Whitney U	1976.500
Wilcoxon W	5216.500
Z	-4.193
Asymp. Sig. (2-tailed)[1]	.000

a. Grouping Variable: Tree type

This test is a non-parametric test, and does not take into consideration the distribution of data. It was performed, as the random tree group did not seem to be normally distributed, so to ensure that the results of the t-test were in fact

[1] P-value

correct. As the distributions of the two graphs were not the same, a difference in the mean ranks was investigated. Prior to carrying out the test, all values are assigned a rank, according to how low or high they are. This way, set maximum and minimum values can be set. This also minimises the effect of any outliers present in the data. The null hypothesis (H_0) was that there is no significant difference between the mean ranks of the two groups of trees. p-value (0.000) < 0.05, therefore null hypothesis, H_0= There is no significant difference between the heights of the two groups, is rejected. The p-value depicts the probability that the test gave a faulty answer. Since it is also <0.001, the difference in the mean ranks is greatly significant.

In what ways do Arundhati Roy and Ken Kesey explore the relationship between voice and silence in *One Flew Over the Cuckoo's Nest* and *The God of Small Things?*

Olivia Brandon - English

For my Extended Essay I chose to explore the literary treatment of Medicine - the subject I am studying next year. After reading a number of texts concerned primarily with mental illness, I decided to write my essay on Ken Kesey's One Flew Over the Cuckoo's Nest and Arundhati Roy's The God of Small Things. The biggest challenge for me was structuring and refining my argument - The God of Small Things is one of my favourite books, and I found myself wanting to talk about everything! Ultimately, I chose to discuss the relationship between elective mutism and the silencing of minority groups in society. Overall, the process was very rewarding - writing my Extended Essay on English Literature has given me a slightly different perspective on Medicine, which I hope to be useful in the years to come.

Supervisor: Mark Beverley

Olivia's Extended Essay is erudite, perceptive and original. Her initial research into literary representations of scientific, medical and psychological practices led to an interest in muteness as a condition of both confinement and liberation. The way in which she argues the validity of this paradox - in reference to both individual fictional characters, as well as the way they stand for cultural norms, practices or historical events, is highly intelligent and hugely compelling. I can think of few essays I have read in quite a number of years that reveal such independence of insight, as well as sophistication and control in analytical style, structure and argument. She was a joy to supervise, and her essay a real pleasure to read.

Contents

Abstract

This essay compares ***the relationship between voice and silence in The God of Small Things and One Flew Over the Cuckoo's Nest,*** examining the process by which silence moves from an expression of the disenfranchised to a means of voice and renewal.

In the first section of the essay I focus on the two authors' presentation of societal voicelessness, examining their shared interest in the experience of the marginalised and the dispossessed. I then explore the link between individual exclusion and the wider displacement of cultural identity presented in both novels. In the second section, I explore the way in which the figurative voicelessness of excluded societal groups is translated into literal voicelessness. I argue that this transition can be seen in the fracture of narrative voice and the presence of two mute characters - Esthappen in *The God of Small Things*, and Chief Bromden in *One Flew Over the Cuckoo's Nest*. By analysing the gradual erosion of Bromden and Estha's voices, I attempt to unpick the origins of their silences. In the essay's third section, I explore the paradox present in both novels' depiction of silence, arguing that silence is both a marker of lost voice and the key to regaining it. I then examine the role of narrative in the re-discovery of voice, demonstrating that in both novels, voice is regained through the imaginative re-creation involved in literary construction.

I conclude that despite originating from different cultures, the key characteristics of the relationship between voice and silence are shared, and that this offers an important reflection on the universal nature of silencing, the psychological origins of muteness, and the prospects of the marginalised for regaining voice.

Introduction

The God of Small Things and *One Flew Over the Cuckoo's Nest* share an interest in the relationship between voice and silence. Both novels repeatedly describe the silencing of specific groups in society, whether they are divorced women, psychiatric patients, or members of India's Dalit caste. Despite being written in different cultures and time periods, there are many similarities between the books in the circumstances that surround this silencing. In both novels, for instance, the loss of individual narratives takes place against a backdrop of lost cultural identity: in Roy's novel we see Indian culture distorted by insidious neo-colonialism, whilst in *One Flew Over the Cuckoo's Nest*, Kesey chronicles the more sudden destruction of Native American culture by corporate exploitation.

In both novels the act of silencing is most intimately expressed through a mute character. In Kesey's novel this character is Chief Bromden; in Roy's, Estha. The immediate origins of their silences are very different - one character is a humiliated Native American haunted by his tribe's destruction, whilst the other is a young Indian man traumatised by sexual abuse and unintended involvement in murder. However, the essential characteristics and complexities of muteness cut across these differences. In both novels muteness is, on one level, presented as an internalisation of societal voicelessness and a means of disengaging from the outside world. However, on another level, it is expressed not as an absence of voice or symbol of passivity, but as a means of maintaining identity and drawing attention to the gaps in what is heard. The paradox inherent in this dualistic presentation of silence is, in both novels, part of a wider focus on the mechanisms not only by which voice is lost, but also by which it is regained. For both authors the rediscovery of voice is ultimately centred on narrative. Through storytelling, characters are able to reverse the process of silencing, allowing them to challenge the prescriptive narratives of their societies and overcome the traumas of the past.

1. Lost voices

The God of Small Things and *One Flew Over the Cuckoo's Nest* both describe the exclusion of certain groups from society, mourning the lost narratives of those who live on the margins and exploring what it means to be left unheard. Kesey's novel is set in an Oregon ward for psychiatric patients. It quickly becomes clear that Kesey regards mental illness not as a medical condition, but as a label given to those who flout social expectations and must therefore be banished from the public domain. Roy's novel, set in Kerala, describes the silencing of multiple groups by caste, patriarchy, adult exploitation, and an indelible colonial legacy. Over the course of the novel these various acts of silencing accumulate with devastating effect, destroying the lives of Roy's seven-year-old characters, Rahel and Esthappen. Roy's focus on dispossession is continuous. From the novel's Untouchables, once expected 'to crawl backwards with a broom, sweeping away their footprints'[1], to the exiled mother who wakes up in 'the strange bed in the strange room in the strange town' with 'only her fear [as] familiar'[2], to the two children haunted by their knowledge that 'when you hurt people, they begin to love you less'[3], her characters are either governed by the threat of exclusion or defined by the experience of it. Likewise, Kesey's psychiatric patients are identified as 'mistakes made in the neighbourhoods and in the schools and in the churches'[4], who can hope only 'to live, if not right in the world of men, at least on the edge of it.'[5]

In both novels we witness the psychological consequences of this failure to conform to prescriptive cultural narratives - narratives that seek to define the shape of a life and invalidate those that do not fit them. In Kesey's novel, Harding, a psychiatric patient, describes these consequences as 'Guilt. Shame. Fear. Self-belittlement', explaining, 'it was the feeling that the great, deadly, pointing forefinger of society was pointing at me - and the great voice of millions chanting,

[1] Roy, A. 1997. *The God of Small Things*. (London, Flamingo). p.73-4.
[2] Ibid, p.161-2
[3] Ibid, p.112
[4] Kesey, K. 1962, *One Flew Over the Cuckoo's Nest.* (London, Methuen & Co Ltd). p.36
[5] Ibid, p.108

'Shame. Shame. Shame.'[6] Here, the menacing, unified chant of society illustrates the drowning out of individual voices and the impossibility of escaping social norms. Likewise, as a divorcee Ammu is haunted by the 'cold knowledge' that life outside social expectations is no life at all: 'life had been lived. She made a mistake. She married the wrong man.'[7]

Invalidated by the dominant narratives of their societies, characters repeatedly experience a loss of voice and identity. In Roy's novel she describes Velutha, a member of the Untouchable caste, as leaving 'no footprints in sand, no ripples in water, no image in mirrors.'[8] Even mirrors - the simplest tools of identity and self-awareness, bear no traces of his presence: he is identified only by absence. Later on, when Velutha is betrayed by the leader of Kerala's communist party, this invisibility is translated into inaudibility: 'Velutha heard his own voice beat back at him as though it hit a wall... The man he was talking to was small and far away, behind a wall of glass.'[9] Here, Roy's simile demonstrates the impenetrability of the social and political barriers that separate Velutha from Touchable society. In Kesey's novel this social exclusion is extended to a detachment from the human race itself: 'all of us in here are rabbits of varying ages and degrees, hippity-hopping through our Walt Disney world.'[10] Here, Kesey's placement of the patients within the realm of children's cartoons evokes the infantilising treatment of mental illness.

In both novels the erosion of individual identity occurs against a backdrop of displaced cultural identity. One of Roy's most powerful depictions of silencing is her account of the Kathakali performances. To accommodate the short attention spans of Western tourists, 'ancient stories [are] collapsed and amputated' and 'six-hour classics... slashed to twenty-minute cameos'.[11] The Kathakali performers, forced by poverty to collude in this commodification of Indian culture, become 'unviable. Unfeasible. Condemned goods'[12], and must '[stop] at the temple to ask the pardon

[6] Kesey, Opp. Cit. p.241
[7] Roy, Opp. Cit., p.38
[8] Ibid, p.216
[9] Ibid, p.287
[10] Kesey, Opp. Cit., p.55
[11] Roy, Opp. Cit., p.127
[12] Ibid, p.230

of their gods. To apologize for corrupting their stories. For encashing their identities. Misappropriating their lives.'[13] Here the language of finance and economics: 'unviable', 'unfeasible', 'condemned goods', and 'encashing', transforms an invaluable means of cultural identification into a marketable product. The loss of the Kathakali stories illustrates not only the alienation of a particular societal group, but also a wider sense of cultural silencing, whereby ancient religious and cultural narratives are gradually excluded from the contemporary domain. As Chacko asserts: 'Our minds have been invaded by a war.... a war that captures dreams and re-dreams them... We belong nowhere.'[14] Here, as in the previous passage, the language of war is employed, but this is not a war with the usual casualties - it is a war whose victims are dreams, stories, minds.

Likewise, the loss of identity experienced by Bromden's father is directly linked to a loss of cultural identity. As government pressure mounts to sell tribal land he asserts, 'What can you pay for the way a man lives?... What can you pay for what a man is? They didn't understand.'[15] As with Roy's Kathakali performers, their identities sold on the tourist market, Bromden's father falls victim to a society willing to sacrifice its own history to corporate greed. If culture and tradition can be seen as the collective voice of a people, then this is another form of silencing - one that is especially brutal for its destruction of not just a person, but a way of life. In this way, both novels link a collective societal voicelessness to the voicelessness of individuals, thereby widening the scope of lost identity and voice.

[13] Ibid, p.229
[14] Ibid, p.53
[15] Kesey, Opp. Cit., p.171

2. The erosion of voice

i.) The motif of silence

In both novels, societal voicelessness is expressed most powerfully through a mute character. Both Bromden and Estha elect to stop speaking, so that a figurative inability to speak becomes a literal one. The two characters share many similarities in the origins of silence. In both novels, individuals who cannot speak are presented as the product of a society that cannot hear. To be mute is to internalise the utter invisibility of life on the periphery of society: as Bromden puts it, 'it wasn't me that started acting deaf; it was people that first started acting like I was too dumb to hear or see or say anything at all.'[16] This sense of imposed voicelessness is extended during Bromden's encounter with government officials: 'the apparatus inside them take[s] the words I just said... when they find the words don't have any place ready-made where they'll fit, the machinery disposes of the words like they weren't even spoken.'[17] Here, the mechanised processing and disposal of language indicates the rejection of narratives incompatible with society's way of seeing, as well as suggesting a breakdown in natural human communication.

Likewise, Estha's inability to speak is an extension of repeated experiences of voicelessness: as Emily Stockdale explains, when the words 'he both hears and speaks fail him', Estha 'releases himself from the constraints' of language.[18] Throughout Roy's novel, Estha's growing voicelessness is signified by the recurrence of the silence motif. For example, in the moments before Estha is abused Roy presents a disturbing image of voice and its erosion. In the cinema auditorium, Estha cannot help but sing along to the *Sound of Music*: his voice is 'clear and true, cutting through the

[16] Ibid, p.163
[17] Ibid, p. 165-6
[18] Stockdale, E. 2008. p.25-6. *Language and the Creation of Characters in Arundhati Roy's The God of Small Things.* [online] Available at: <http://libres.uncg.edu/ir/uncw/f/stockdalee2008-1.pdf> [Accessed 26 August 2014]

fan-whirring, peanut-crunching darkness.'[19] And yet, in a show of adult hostility, Estha is symbolically silenced by members of the audience with 'hissing mouths' and 'teeth like sharks'[20]: here, Roy forecasts the predatory malice of adults through animalistic vocabulary that sharply contrasts Estha's innocence and purity. In the aftermath of trauma, silence is internalised and translated into a breakdown in language. Suddenly unable to verbalise his emotions, Estha dissociates from them by describing not his own helplessness but that of the 'little banana bubbles drowning deep in jam and nobody to help them.'[21] In this way, language loses its capacity to define and thus work through trauma.

An element of Estha's muteness that is absent from Kesey's novel involves a fear of language and its consequences. In the police station Estha condemns Velutha to death in a single word. His resulting sense of betrayal is so overwhelming that he becomes identified with Judas through the novel's biblical parallelism. Velutha - a carpenter, is tortured and then executed by government officials, and, like Christ, identified and thus condemned by one supposedly most loyal to him. And yet, the story is different, Estha is coerced into betrayal - and so he becomes enveloped in a new form of voicelessness. Forced into a lie, he loses his voice to Baby Kochamma's narrative of revenge; furthermore, he is introduced to the destructive power of language, so that later in the novel he convinces himself that in predicting the permanence of his separation from Ammu he has made it a reality: 'Because he was the one that had *said* it. *But Ammu that will be never!'*.[22] Thus, silence becomes rooted not only in the internalisation of voicelessness but also in a fear of language's capacity for destruction.

[19] Roy, Ibid, p.100
[20] Ibid, p.100
[21] Ibid, p.194
[22] Ibid, p.325

ii.) Trauma and the fracture of narrative voice

In both novels the breakdown of voice is also communicated through the use of fragmented narratives, which indicate the disconnection between characters and their voices and the debilitating effects of trauma on the mind. Fragmentation takes a number of forms in both novels. Both authors apply structural fragmentation in the shape of non-linear narratives and alternating storylines in order to distort regular chronology and thus mimic the psychological impact of trauma. Both narratives are also scattered with recurrent images, smells, and sounds, often left unexplained until far later in the novel. For example, in Roy's novel the smell of 'old roses on a breeze'[23] is repeatedly referred to, but only connected to Velutha's blood in the novel's final pages. The insistent return of the image mirrors the impact of trauma, enabling the reader to experience second-hand the repressed memories and fractured thoughts of Roy's characters. At other points in Roy's novel, memories are transmitted across generations, fracturing the linear motion of the narrative. For example, the ghost of Pappachi's moth appears persistently to '[torment] him and his children and his children's children'[24] - an emblem of familial trauma that seems to speak for all the small, seemingly negligible tragedies that coalesce over Roy's novel to drive its central calamity. Likewise, in Kesey's novel trauma is broken down and encoded in fragments of memory: we hear the 'slap of a salmon's tail on water'[25] during one of Bromden's hallucinations, but only able to appreciate its origins much later on.

In Kesey's novel the dislocating effect of trauma is revealed in the contrast between Bromden's language in the past and the present. Bromden's flashbacks are related largely through imagery of the natural world: his father was born 'smack into' his name 'the way a calf drops out in a spread blanket'[26] and the scaffolding on the falls is revealed as a natural organism 'that has been

[23] Ibid, p.32
[24] Ibid, p.49
[25] Kesey, Opp. Cit., p.72
[26] Ibid, p.224

growing and branching out among the rocks... for hundreds of years.'[27] This use of natural imagery is implicitly connected to the oral traditions of Bromden's tribe and their emphasis on man's connection with nature. However, throughout Kesey's novel Bromden's metaphors are largely mechanical: the red capsules patients are given are really 'miniature electronic elements'[28], and Bromden 'can smell the machinery inside [the nurse] the way you smell a motor pulling too big a load'[29]. Through the dominance of this mechanical imagery, Kesey signals the dislocating effect of Bromden's separation from the natural world. In this way, evidence of trauma in the narrative voice demonstrates the struggle characters face in accounting for a present dominated by the past.

3. Regaining voice

i.) Silence as voice

Neither character's muteness can be interpreted solely within the frame of loss and passivity. Elaine Ware sees Bromden's silence as a 'survival technique'[30]. Through his 'deaf-and-dumb act'[31], Bromden is able to '[listen] to secrets meant only for their ears'[32], giving him a unique perspective into staff decisions. Estha's muteness can also be seen as a protection mechanism. Roy asserts of his silence, 'it reached out of his head and enfolded him in its swampy arms. It rocked him to the rhythm of an ancient, foetal heartbeat.'[33] Here, language and imagery associated with the maternal - 'enfolded him in its swampy arms', 'rocked him', 'foetal' - evokes the soothing nature of silence, alluding to it as a means of withdrawing to a state of womb-like comfort and dependency. This

[27] Ibid, p.166
[28] Ibid, p.32
[29] Ibid, p.10
[30] Ware, E. 1986. p.2 *The Vanishing American: Identity Crisis in Ken Kesey's One Flew Over the Cuckoo's Nest*. [online] Available at: <http://apenglish1011.weebly.com/uploads/2/4/3/5/2435985/the_vanishing_american.pdf> [Accessed 26 August 2014]
[31] Kesey, Opp. Cit., p.24
[32] Ibid, p.117
[33] Roy, Ibid, p.11

desire to return to the early stages of life is furthered by a sense of the primitive in the words 'swampy' and 'ancient'.

In both novels, silence not only acts as a survival mechanism, but is itself a form of voice. Through muteness, Bromden and Estha highlight all that remains unrecognised in the dominant narratives of their societies, whether it is Native American culture, mental illness, or sexual abuse. Silence becomes as powerful a form of communication as its opposite: in Kesey's novel this paradox is evident in a 'voiceless' character whose voice is the only one we hear. Early in the novel Bromden asserts, 'they think I'm deaf and dumb... I'm cagey enough to fool them that much.'[34] Silence is more than the internalisation of voicelessness or even a need for passive survival: it is also a form of resistance that allows him to make fools out of those who have humiliated him. Similarly, whilst Roy denies that Estha's silence is 'intrusive' or 'accusing'[35], its role as a form of protest can be demonstrated by Comrade Pillai's reaction to it: '[he] would slap himself all over to get his circulation going. He couldn't tell whether Estha recognised him after all those years.'[36] There is also a sense that in the context of Estha's past, in which speech has been insisted upon to devastating effects, silence becomes not only an act of resistance, but also an assertion of freedom. Thus, whilst in both novels silence remains a marker of fractured identity and powerlessness, on another level it serves to counter the very voicelessness it represents. In this there is a central paradox: the loss of voice is also the key to regaining it.

[34] Kesey, Opp. Cit., p.9
[35] Roy, Opp. Cit., p.10
[36] Ibid, p.14

ii.) Narrative as the vehicle for rediscovery

In both novels, the rediscovery of voice occurs largely through narrative. Storytelling becomes a means of working through trauma, of giving structure to experience, and crucially, of reclaiming authority over one's own narrative. In Kesey's novel, the rediscovery of voice is made obvious by Bromden's return to speech and escape from the ward. In Roy's it is more complex: Estha does not, like Bromden, learn to speak again. However, in employing a diversity of perspectives Roy gradually retrieves her characters' lost narratives. No individual story or personal loss goes unheard: this plurality of voices is introduced in Roy's epigraph, which states, 'never again will a single story be told as if it's the only one.'[37] In this, Roy rejects any dominant narrative that seeks to impose itself as the single, authoritative voice of her characters' history, instead bringing to light the many different perspectives that make up a story. In addition to this, we witness a central recovery of identity in Roy's two main characters. Joya Chakravarty states that the novel 'ends in a feeling of loneliness and despair'[38]. I would argue that this is far from the case, and that through the novel the twins achieve a form of resolution that suggests a coming to terms with the past and the reconstruction of identity. This can be seen in the way Rahel's perception of her mother changes over the novel. Initially Roy presents Rahel's disgust at Ammu: 'She thought of the phlegm and nearly retched. She hated her mother then. *Hated* her.'[39] However, as Ammu's own perspective is explored, Rahel gradually reaches an understanding of her actions, and with understanding, reconciliation. In the novel's final stages Rahel remembers a dream: 'a fat man, faceless, kneeling beside a woman's corpse. Hacking its hair off. Breaking every bone in its body... A pianist killing

[37] Ibid, epigraph
[38] Charkravarty, J. 2003. p.95. *Indian Writing in English: Perspectives*. (New Delhi, Atlantic Publishers and Distributors). p.95 [online] Available at: <http://books.google.co.uk/books?id=QibxKNBmCaIC&printsec=frontcover&dq=Indian+Writing+in+English:+Perspectives&hl=en&sa=X&ei=UUb8U5ikDs7c4QTQq4G4DA&ved=0CCgQ6AEwAQ#v=onepage&q=Indian%20Writing%20in%20English%3A%20Perspectives&f=false> [Accessed 26 August 2014]
[39] Roy, Opp. Cit., p.161

the piano keys.'[40] Here, for the first time, Rahel recognises her mother as a victim. The slow, sadistic breaking of bones evokes Ammu's gradual wearing down by 'faceless' patriarchy, whilst the 'hacking' of Ammu's hair suggests her unwarranted treatment as a prostitute. The comparison of Ammu to a broken piano condenses her treatment in a vivid image of silencing. Roy concludes, 'And Rahel... loved them both. The player and the piano. The killer and the corpse.'[41] In this there is a sense of regained familial identity, with Rahel recognising the futility of choosing sides, and the possibility that those she loves have been both victims and perpetrators. Thus, Roy's novel is a story of restored voice and identity as much as Kesey's is.

In both novels, therefore, narrative not only communicates the voices of the voiceless, but is itself a means of re-discovery. In *Trauma Fiction* Anne Whitehead describes writing as an 'inherently curative process'.[42] This idea is evident in both books. Through narrative, Bromden and Rahel give structure and meaning to their experiences, threading together various events and time periods to create a coherent account of who they are and why. Traumatic experiences are gradually confronted, and with each experience explored, they come a step closer to understanding the past. Elaine Ware states that it is 'only through the support of McMurphy' that Bromden 'regain[s] his strength' and 'self-confidence.'[43] Although McMurphy unquestionably acts as a catalyst for re-discovery, it is no coincidence that Bromden's return to speech occurs immediately after he addresses the loss of Columbia Gorge. In narrating the experience, Bromden confronts his most painful source of trauma, thereby uncovering the origins of his silence: 'It was people that first started acting like I was too dumb to hear or see or say anything at all'.[44] Likewise, Roy's narrators

[40] Ibid, p.225
[41] Ibid, p.225
[42] Whitehead, A. p.87. *Trauma Fiction.* (Edinburgh, Edinburgh University Press Ltd). p.87 [online] Available at: <http://books.google.co.uk/books?id=6bV7EGoYzRIC&printsec=frontcover&dq=Trauma+Fiction+Anne+Whitehead&hl=en&sa=X&ei=_Eb8U9CZGtL14QSViID4Cw&ved=0CCsQ6AEwAA#v=onepage&q=Trauma%20Fiction%20Anne%20Whitehead&f=false> [Accessed 26 August 2014]
[43] Ware, Opp. Cit., p.4
[44] Kesey, Opp. Cit., p.163

have a fundamental need to organise and thereby understand the past. In a way that almost seems to constitute the novel's introduction, Roy asserts:

> 'Things can change in a day... when they do, those few dozen hours, like the salvaged remains of a burned house - the charred clock, the singed photograph, the scorched furniture - must be resurrected from the ruins and examined. Preserved. Accounted for.'[45]

Here there is a sense of archaeology in Roy's meticulous study of the past, suggesting the sifting through of multiple layers of memory in search of a legitimate identity for her two surviving protagonists.

The 'curative' nature of storytelling can also be demonstrated by the strong sense of liberation it brings. With narrative comes an assertion of authority in the right to govern one's own story, and thereby reject the dominant narratives of family, society, and culture. This sense of liberation can be demonstrated by Bromden's assertion, 'I been silent so long now it's gonna roar out of me like floodwaters.'[46] Throughout Kesey's novel, water is associated with freedom. For example, the symbolic shattering of Nurse Ratched's window is compared to 'water splashing'[47], and on throwing the control panel through the screen, 'the glass splashe[s] out in the moon, like a bright cold water baptizing the sleeping earth.'[48] Significantly, it is the men's trip to the ocean that ultimately restores freedom and identity. Thus, in addition to signifying the sheer force of speech, Kesey's water imagery suggests the liberating effect of narrative. The simile also links Bromden's narrative to the restoration of his cultural heritage: Kesey's association between water and identity is partly rooted in the falls once owned by Bromden's tribe. Interestingly, Estha's reengagement with the world is also described as a flood of water:

[45] Roy, Opp. Cit., p. 32
[46] Kesey, Opp. Cit., p.12
[47] Ibid, p.155
[48] Ibid, p.254

> 'The world, locked out for years, suddenly flooded in, and now Estha couldn't hear himself for the noise... A dam had burst and savage waters swept everything up in a swirling.'[49]

Here, as for Bromden, we witness the sudden return of voice that initiates Roy's novel: through storytelling, repressed thoughts and memories are finally allowed to surface, and the barrier imposed between Estha and his world quickly begins to fragment. Roy's use of dam imagery is, like Kesey's, very culturally relevant. For Roy, dams are heavily associated with the voiceless: her essay 'The Greater Common Good' chronicles the mass displacement of Indians by 3,200 dams - dams whose effects have been left unrecorded in a way that has 'strangl[ed] stories'[50] and 'squashed' people 'like bugs'[51]. Thus, in both novels, the metaphorical collapse of dams conveys the breakthrough of voice enabled by narrative. In this way, Roy and Kesey present the power of narrative to reverse the process of silencing.

[49] Roy, Opp. Cit., p.15
[50] Roy, A. 1999. *The Cost of Living.* (London, Flamingo). p.72
[51] Ibid, p.98

Conclusion

In both novels the relationship between voice and silence is multifaceted, indeed paradoxical. However, crucially, the key characteristics of this relationship are shared. Both authors employ disorder and fragmentation as a means of reconstruction. Both demonstrate the devastating impact of exclusion, and yet celebrate the differences that give rise to it. Both describe the lost narratives and hidden lives of those relegated to the margins of society, whilst simultaneously denying through their novels that conceptually, the 'voiceless' exist. For authors of completely different cultures and backgrounds, with completely different political causes and concerns, the origins of silence and the extent of societal voicelessness are remarkably similar. However, in the end, significant problems remain. Although characters may regain their voices within the books, the question of providing representation outside them is not as simple. Much of Kesey's novel, for example, is structured around a Native American's rediscovery of voice, and yet, this rediscovery must be questioned when in many ways Kesey simply further envelops Bromden in the prescriptive narratives of alcoholism, white oppression, and intergenerational trauma. Equally, Roy repeatedly describes the voicelessness of Untouchables, but it could be argued that romanticising and deifying them is just another way of silencing them. Both novels ultimately raise the question of whether it is ever possible to give a voice to the 'voiceless' - authors can try to give people a voice, they may even try to give them their own voice, but in the end, it is always their own.

Bibliography

Charkravarty, J. 2003. *Indian Writing in English: Perspectives*. (New Delhi, Atlantic Publishers and Distributors). p.95 [online] Available at: <http://books.google.co.uk/books?id=QibxKNBmCaIC&printsec=frontcover&dq=Indian+Writing+in+English:+Perspectives&hl=en&sa=X&ei=UUb8U5ikDs7c4QTQq4G4DA&ved=0CCgQ6AEwAQ#v=onepage&q=Indian%20Writing%20in%20English%3A%20Perspectives&f=false> [Accessed 26 August 2014]

Kesey, K. 1962, *One Flew Over the Cuckoo's Nest.* (London, Methuen & Co Ltd).

Roy, A. 1999. *The Cost of Living.* (London, Flamingo).

Roy, A. 1997. *The God of Small Things*. (London, Flamingo).

Stockdale, E. 2008. 'Language and the Creation of Characters in Arundhati Roy's The God of Small Things'. [online] Available at: <http://libres.uncg.edu/ir/uncw/f/stockdalee2008-1.pdf> [Accessed 26 August 2014]

Ware, E. 1986. 'The Vanishing American: Identity Crisis in Ken Kesey's One Flew Over the Cuckoo's Nest'. [online] Available at: <http://apenglish1011.weebly.com/uploads/2/4/3/5/2435985/the_vanishing_american.pdf> [Accessed 26 August 2014]

Whitehead, A. *Trauma Fiction.* (Edinburgh, Edinburgh University Press Ltd). p87 [online] Available at: <http://books.google.co.uk/books?id=6bV7EGoYzRIC&printsec=frontcover&dq=Trauma+Fiction+Anne+Whitehead&hl=en&sa=X&ei=_Eb8U9CZGtL14QSViID4Cw&ved=0CCsQ6AEwAA#v=onepage&q=Trauma%20Fiction%20Anne%20Whitehead&f=false> [Accessed 26 August 2014]

A vector approach to 3D reconstruction from multiple views

Loris Gliner - Mathematics

For my Mathematics Extended Essay I produced a report entitled "A Vector Approach to 3D Reconstruction from Multiple Views". It develops an algorithm to build a 3D model of an object from multiple 2D images displaying it. This wasn't without challenges: Firstly, I came across complex cutting edge literature so I set up a simple base case and gradually extended it to reach the ultimate objective. Then, having grasped the concept, I had to find a way to express myself in a clear and concise way (at least I tried…). Nonetheless, this long yet exciting journey, was filled with many satisfactory "AHA" moments, particularly when resolving the issues I encountered!

Supervisor: Guy Howden

Loris undertook a significant amount of personal research to complete his essay, inquiring into new techniques and methods hitherto un-encountered. Inspired by a talk in Maths Society on Image Reconstruction, Loris developed a genuine interest in the intricate Mathematics involved. Upon exploring the original concepts a more creative approach had to be undertaken in order to simplify a complex problem – an approach stemming from his natural curiosity and desire to understand the nature of the problem. Loris' ability to elucidate mathematical concepts expressing his methods with clarity enabled him to produce a computational model to reconstruct a 3D image using Mathematica and produce an essay which evidences his own personal journey.

Contents

Abstract

In this Mathematics Extended Essay, the research question is to develop **"A vector approach to 3D reconstruction from multiple views"**.

In order to develop this approach, I have first defined the Pinhole Camera model, which is a model describing the mapping from a set of 3D coordinates representing an object in 3D space to a set of 2D coordinates on a plane, representing an image of this object.

From this model, the next step of this investigation is to derive a reconstruction algorithm given two views. In addition, an example is shown.

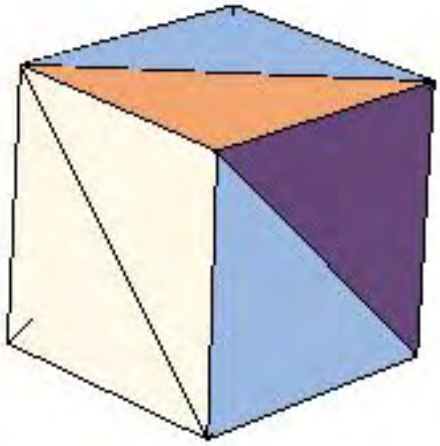
Then, I had the opportunity to participate to the Wolfram Mathematica Summer Camp in July 2014 where I conceived and programmed a demonstration program illustrating and verifying this vector approach to 3D reconstruction from multiple views.

Throughout this essay a vector notation is used in order to clarify the mathematical processes. As a result of this notation, coordinates are referred to as position vectors.

The outcome of this investigation is a solution to the research question given **two views** taken by **calibrated cameras** with **manually matched corresponding points**. (Cameras are considered calibrated when information about the positions of the cameras relative to each other are known).

Only visible corresponding points were reconstructed, therefore, with only two views, some coordinates of the object used in the example were missing.

Further improvements to this research include automatic detection of corresponding points, camera self- calibration and introducing more than two views.



2 Introduction

3D reconstruction from multiple 2D views is nowadays an active area of research in the field of computer vision as it has a wide range of practical applications including medical imaging and also the film and television industries (this is the technology used to monitor if the ball is in or out on a tennis court in international tournaments for example).
Furthermore it is a very interesting area of applied mathematics using 3D coordinates geometry.

In this essay it will be attempted to develop **"A vector approach to 3D reconstruction from multiple views"** expanding on the knowledge acquired during the International Baccalaureate Mathematics Higher Level course.

This essay will cover the case of calibrated cameras (some information relative to the cameras used to take the photographs and their position will be known) and manually selected corresponding points between two images.
First the model of the reconstruction will be explained and then the algorithm of the reconstruction will be derived.

2.1 Notations in the essay

A scalar, (one dimensional quantity), will be represented by a lower case letter: $s \in R$.

The axes in a 3D space scene will be X, Y, Z.

The axes in a 2D space scene will be x, y.

A position vector will be represented by a bold letter with a right arrow on top of it.

- In the case of a 3D position vector it will be an upper case letter: $\overrightarrow{\mathbf{P}} = \begin{pmatrix} X_{\mathbf{P}} \\ Y_{\mathbf{P}} \\ Z_{\mathbf{P}} \end{pmatrix}$.
- In the case of a 2D position vector it will be a lower case letter: $\overrightarrow{\mathbf{p}} = \begin{pmatrix} x_{\mathbf{p}} \\ y_{\mathbf{p}} \end{pmatrix}$.

A direction vector will be represented by a two bold letters with a right arrow on top of them.

- In the case of a 3D direction vector it will be upper case letters: $\overrightarrow{\mathbf{QR}} = \begin{pmatrix} X_{\mathbf{QR}} \\ Y_{\mathbf{QR}} \\ Z_{\mathbf{QR}} \end{pmatrix}$.
- In the case of a 2D position vector it will be lower case letters: $\overrightarrow{\mathbf{qr}} = \begin{pmatrix} x_{\mathbf{qr}} \\ y_{\mathbf{qr}} \end{pmatrix}$.

A unit vector will be represented by a bold letter with a circumflex on top: $\widehat{\mathbf{U}} = \dfrac{\overrightarrow{\mathbf{U}}}{|\overrightarrow{\mathbf{U}}|}$.

A vector equation of a plane will be represented by the symbol Π followed by the name of the plane (in lower case): $\Pi p = \overrightarrow{\mathbf{P}} + \overrightarrow{\mathbf{QR}} + \overrightarrow{\mathbf{ST}}$.

3 Definition of the Pinhole Camera model

A pinhole camera model is a model describing the mapping from the set of 3D coordinates representing an object in 3D space to the set of 2D coordinates on a plane representing an image of this object.

Note: In the pinhole camera model, all the light rays of the set of 3D coordinates representing an object in 3D space converge in a position vector, called the camera centre, $\overrightarrow{\mathbf{C}}$, as shown in figure 3.1.

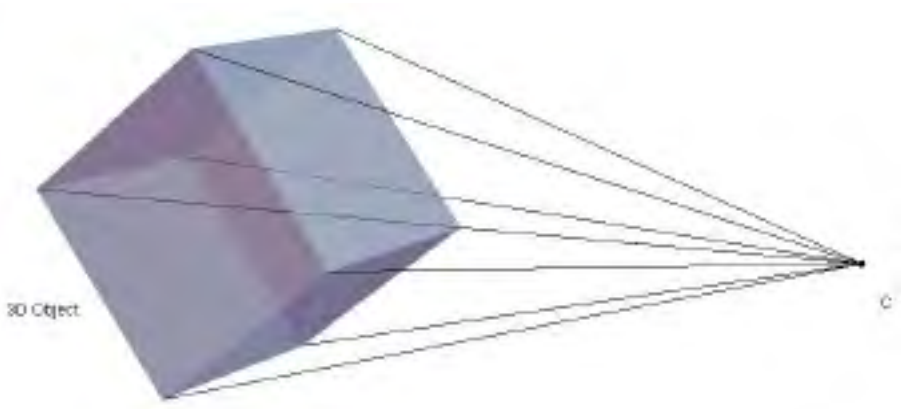


Fig. 3.1. All the rays of the 3D object converge to the camera center, $\overrightarrow{\boldsymbol{C}}$.

Definition:
Consider the position vector $\overrightarrow{\mathbf{x}}_i = \begin{pmatrix} x_{\mathbf{x}_i} \\ y_{\mathbf{x}_i} \end{pmatrix}$ on the 2D image plane $\Pi i = \overrightarrow{\mathbf{P}} + \overrightarrow{\mathbf{QR}} + \overrightarrow{\mathbf{ST}}$ in 3D space, which is the image of the point $\overrightarrow{\mathbf{X}} = \begin{pmatrix} X_{\mathbf{X}} \\ Y_{\mathbf{X}} \\ Z_{\mathbf{X}} \end{pmatrix}$ in 3D space.

Under the pinhole camera model, the point $\overrightarrow{\mathbf{x}}_i$ is defined as the intersection between the direction vector $\overrightarrow{\mathbf{XC}}$, (where $\overrightarrow{\mathbf{C}}$ is the camera centre), and the image plane Πi.
This is shown in figure 3.2.

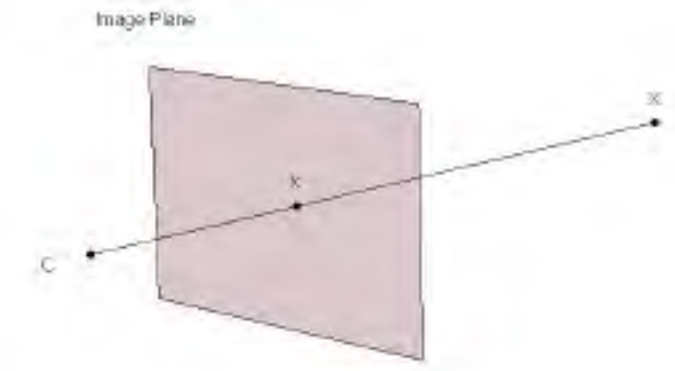


Fig. 3.2. The pinhole camera model.

This is the image created by a pinhole camera, where there is a slit instead of a lens. In a conventional camera, this is not exactly true due to geometric distortions or blurring of unfocused objects.
However, in this essay we will consider that conventional cameras follow the pinhole camera model as the error is relatively small.

Note: When using a conventional camera the distance between $\overrightarrow{\mathbf{C}}$, the camera centre, and the image plane is equal to the focal length of the camera.

Note 2: When using a conventional camera the image taken by the camera is represented on the image plane by a rectangle ($\overrightarrow{\mathbf{O}_1}\overrightarrow{\mathbf{L}}\overrightarrow{\mathbf{M}}\overrightarrow{\mathbf{N}}$) with a size equal to the image capture size and, $\overrightarrow{\mathbf{C}}$, the camera centre, is on the vector perpendicular to the plane and passing through the center of the rectangle.

This model gives us an important result: When taking an image of a 3D object, a dimension (depth) is lost, as shown in figure 3.3.

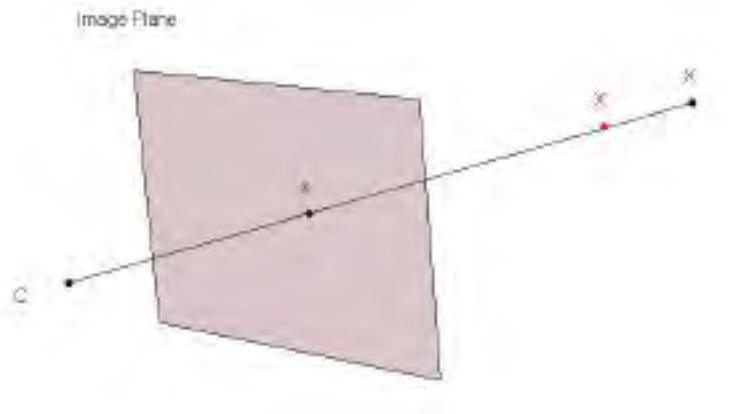


Fig.3.3. Two different points in 3D space ($\overrightarrow{\boldsymbol{X}}$ and $\overrightarrow{\boldsymbol{X'}}$) will have the same point on the image ($\overrightarrow{\boldsymbol{O}_i}$). This shows the important result that depth is lost in a 2D image of a 3D object.

However, by adding a second image of the 3D object, information about the depth can be retrieved. This is shown by figure 3.4.

Fig.3.4. If we are using only image 1, depth information is lost because every point on the line $\overrightarrow{\boldsymbol{CX}}$ (such as $\overrightarrow{\boldsymbol{Y}}$) projects to the same point on the image 1. However, on image 2, different points on the line $\overrightarrow{\boldsymbol{CX}}$ projects to different points.

This is the principle that will be used to reconstruct the 3D object: Knowing the position vectors $\overrightarrow{\mathbf{C}}$, $\overrightarrow{\mathbf{C'}}$, $\overrightarrow{\mathbf{X}}_i$ and $\overrightarrow{\mathbf{X}}_{i'}$, (where the 3D coordinates of $\overrightarrow{\mathbf{x}}_i$ and $\overrightarrow{\mathbf{x}}_{i'}$ are respectively $\overrightarrow{\mathbf{X}}_i$ and $\overrightarrow{\mathbf{X}}_{i'}$), it is then possible to triangulate to the correct 3D coordinate, $\overrightarrow{\mathbf{C}}$.
This is done by finding the position vector at the intersection of the lines ($\overrightarrow{\mathbf{CX}_i}$) and ($\overrightarrow{\mathbf{CX}_{i'}}$). Indeed,

$$\overrightarrow{\mathbf{X}} = \overrightarrow{\mathbf{C}} + k \times \overrightarrow{\mathbf{CX}_i}, \tag{1}$$

Similarly

$$\overrightarrow{\mathbf{X}} = \overrightarrow{\mathbf{C'}} + k' \times \overrightarrow{\mathbf{C'X}_{i'}}, \tag{2}$$

where k and k' are the solutions to the equation $\overrightarrow{\mathbf{C}} + k \times \overrightarrow{\mathbf{CX}_i} = \overrightarrow{\mathbf{C'}} + k' \times \overrightarrow{\mathbf{C'X}_{i'}}$.

4 The Reconstruction Algorithm for 2 Views

From the pinhole camera model, it is possible to reconstruct a 3D model from two 2D images. Furthermore, in order to find the 3D coordinates of the object we need to know the position of $\overrightarrow{\mathbf{C}}$ (the camera centrer of image 1), $\overrightarrow{\mathbf{C'}}$,(the camera centrer of image 2) and the position of $\overrightarrow{\mathbf{x}}_i$ and $\overrightarrow{\mathbf{x}}_{i'}$ (two corresponding points on image 1 and image 2).

Note: all the points need to be in the same coordinate system. This means that the position of the images (or camera used to take the pictures) relative to each other must be known.

Therefore, the input to the reconstruction algorithm must be the following:

- 2 images of a 3D object
- The position of the camera centres $\overrightarrow{\mathbf{C}}$ and $\overrightarrow{\mathbf{C'}}$ relative to an arbitrary placed origin, $\overrightarrow{\mathbf{O}_2}$, in a 3D space scene with 3 mutually perpendicular axes: X, Y, Z.
- For each image plane: Two direction vectors parallel to the image plane and where one of the direction vector is also parallel to x axis of the image plane and going in the same direction as the x axis of the image plane, were the x axis of the image plane contains the corners down to the left and to the right of the image it represents and were the origin of the x axis of the image plane is the corner down to the left of the image it represents.
- The focal length of the camera used to take the images
- The image capture size of the camera used to take the images

The two images should have as much corresponding points as possible as only corresponding points can be reconstructed, non visible corresponding points of a 3D object won't be reconstructed.
In addition, the images should be taken with no zoom as otherwise the focal length of the camera changes (unless the new focal length is known when zooming).

4.1 Summary of the Algorithm

Step 1: Finding the equation of the image planes.

Step 2: Finding the position of the x and y axes and the position of the origin of each image plane.

Step 3: Finding the position of $\overrightarrow{\mathbf{X}_i}$ and $\overrightarrow{\mathbf{X}_{i'}}$ from $\overrightarrow{\mathbf{x}_i}$ and $\overrightarrow{\mathbf{x}_{i'}}$ respectively, (where $\overrightarrow{\mathbf{x}_i}$ and $\overrightarrow{\mathbf{x}_{i'}}$ are corresponding position vectors on the 2D images), to reconstruct $\overrightarrow{\mathbf{X}}$.

4.2 Algorithm

Let i be the image 1 and i' be the image 2.
Let $\overrightarrow{\mathbf{C}}$ be the camera centre of image 1
Let $\overrightarrow{\mathbf{C'}}$ be the camera centre of image 2
Let $\overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{ST}}$ be the two direction vectors parallel to the plane of image 1, where $\overrightarrow{\mathbf{QR}}$ is parallel to and in the same direction as the x axis of the plane of image 1.
Let $\overrightarrow{\mathbf{Q'R'}}$ and $\overrightarrow{\mathbf{S'T'}}$ be the two direction vectors parallel to the plane of image 1, where $\overrightarrow{\mathbf{Q'R'}}$ is parallel to and in the same direction as the x axis of the plane of image 2.
Let f be the focal length of the camera
Let $d = w \times h$ be the image capture size of the images taken by the camera

Step 1: First of all, we need to find the equation of the plane of image 1, Πi, and of the plane of image 2, Πi'.

Knowing that the two vectors $\overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{ST}}$ are parallel to Πi, let Πi contain point the position vector $\overrightarrow{\mathbf{P}}$. Thus we can write

$$\Pi i = \overrightarrow{\mathbf{P}} + \lambda\overrightarrow{\mathbf{QR}} + \mu\overrightarrow{\mathbf{ST}}. \tag{3}$$

Therefore, to find the equation of Πi, we need to find a possible value for $\overrightarrow{\mathbf{P}}$.

Furthermore, knowing that the distance between $\overrightarrow{\mathbf{C}}$ and Πi is f, one possible value of $\overrightarrow{\mathbf{P}}$ is such that:

- $\overrightarrow{\mathbf{P}}$ is at a distance of f to $\overrightarrow{\mathbf{C}}$,
- $\overrightarrow{\mathbf{P}}$ is on a vector perpendicular to the image plane Πi.

Let $\overrightarrow{\boldsymbol{\alpha}}$ be a direction vector perpendicular to the plane Πi, hence we can write

$$\overrightarrow{\mathbf{P}} = \overrightarrow{\mathbf{C}} + k\overrightarrow{\boldsymbol{\alpha}} \tag{4}$$

such that $|\mathrm{k}\overrightarrow{\boldsymbol{\alpha}}| = f$, that k is a scalar and where $\overrightarrow{\boldsymbol{\alpha}}$ can be written as the cross product of $\overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{ST}}$.*

As $|\mathrm{k}\overrightarrow{\boldsymbol{\alpha}}| = f$ we can write

$$k = \frac{f}{|\overrightarrow{\boldsymbol{\alpha}}|}, \tag{5}$$

$$\therefore\ \overrightarrow{\mathbf{P}} = \overrightarrow{\mathbf{C}} + \frac{f}{|\overrightarrow{\boldsymbol{\alpha}}|} \times \overrightarrow{\boldsymbol{\alpha}}, \tag{6}$$

$$\boxed{\therefore\ \Pi i = (\overrightarrow{\mathbf{C}} + \widehat{\boldsymbol{\alpha}} \times f\,) + \lambda\overrightarrow{\mathbf{QR}} + \mu\overrightarrow{\mathbf{ST}},} \tag{7}$$

where $\overrightarrow{\boldsymbol{\alpha}}$ can be written as the cross product of $\overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{ST}}$.*

The same process can be repeated to find the equation of the image plane Πi', replacing vector $\overrightarrow{\boldsymbol{\alpha}}$ with $\overrightarrow{\boldsymbol{\alpha}}$' and $\overrightarrow{\mathbf{P}}$ with $\overrightarrow{\mathbf{P'}}$,

$$\boxed{\therefore\ \Pi i' = (\overrightarrow{\mathbf{C'}} + \widehat{\boldsymbol{\alpha'}} \times f\,) + \lambda\overrightarrow{\mathbf{Q'R'}} + \mu\overrightarrow{\mathbf{S'T'}},} \tag{8}$$

where $\overrightarrow{\boldsymbol{\alpha}}$' can be written as the cross product of $\overrightarrow{\mathbf{Q'R'}}$ and $\overrightarrow{\mathbf{S'T'}}$.†

Step 2: Given that the x axis of the image plane of an image contains the corners down to the left and to the right of the image it represents, and that the origin of the x axis of the image plane of an image is the corner down to the left of the image it represents, let the image planes, Πi and Πi', contain the rectangles $\overrightarrow{\mathbf{O}_1}\overrightarrow{\mathbf{L}}\overrightarrow{\mathbf{M}}\overrightarrow{\mathbf{N}}$ and $\overrightarrow{\mathbf{O'}_1}\overrightarrow{\mathbf{L'}}\overrightarrow{\mathbf{M'}}\overrightarrow{\mathbf{N'}}$ which respectively represent i and i'.
Note: $\overrightarrow{\mathbf{O}_1}$ and $\overrightarrow{\mathbf{O'}_1}$ respectively represent the corner down to the left of the images.
Note 2: $\overrightarrow{\mathbf{L}}$ and $\overrightarrow{\mathbf{L'}}$ respectively represent the corner down to the right of the images.
Therefore $\overrightarrow{\mathbf{O}_1}$ and $\overrightarrow{\mathbf{O'}_1}$ are respectively the origin of the x and y axes of Πi and Πi'. In addition $\overrightarrow{\mathbf{O}_1\mathbf{L}}$ and $\overrightarrow{\mathbf{O'}_1\mathbf{L'}}$ are respectively on the x axis of Πi and Πi'. So $\overrightarrow{\mathbf{O}_1\mathbf{N}}$ and $\overrightarrow{\mathbf{O'}_1\mathbf{N'}}$ are respectively on the y axis of Πi and Πi'.

The next step of this algorithm is to find the position of the origin $\overrightarrow{\mathbf{O}_1}$ on Πi , the position of the origin $\overrightarrow{\mathbf{O'}_1}$ on Πi' and the equations of the x and y axes of each image plane.

Knowing that $\overrightarrow{\mathbf{C}}$ is on a vector perpendicular to the image plane Πi and passing through the center of the rectangle $\overrightarrow{\mathbf{O}_1}\overrightarrow{\mathbf{L}}\overrightarrow{\mathbf{M}}\overrightarrow{\mathbf{N}}$ and knowing from step 1 that $\overrightarrow{\mathbf{P}}$ lies on the image plane Πi and lies on a vector perpendicular to the image plane Πi and through $\overrightarrow{\mathbf{C}}$ we can deduce that $\overrightarrow{\mathbf{P}}$ is the center of the rectangle $\overrightarrow{\mathbf{O}_1}\overrightarrow{\mathbf{L}}\overrightarrow{\mathbf{M}}\overrightarrow{\mathbf{N}}$.

*When looking at the vector $\overrightarrow{\boldsymbol{\alpha}}$ from a point of view where $\overrightarrow{\mathbf{C}}$ is to the left and $\overrightarrow{\mathbf{C'}}$ is to the right, $\overrightarrow{\boldsymbol{\alpha}}$ must go forwards.
†When looking at the vector $\overrightarrow{\boldsymbol{\alpha'}}$ from a point of view where $\overrightarrow{\mathbf{C}}$ is to the left and $\overrightarrow{\mathbf{C'}}$ is to the right, $\overrightarrow{\boldsymbol{\alpha'}}$ must go forwards.

As a result and given that $d = w \times h$ we can say that

$$\overrightarrow{\mathbf{O}_1} = \overrightarrow{\mathbf{P}} - k\overrightarrow{\mathbf{O}_1\mathbf{L}} - k'\overrightarrow{\mathbf{O}_1\mathbf{N}}, \tag{9}$$

such that $|\mathrm{k}\overrightarrow{\mathbf{O}_1\mathbf{L}}| = \frac{w}{2}$ and that $|\mathrm{k'}\overrightarrow{\mathbf{O}_1\mathbf{N}}| = \frac{h}{2}$ and that k and k' are scalars.

As $|\mathrm{k}\overrightarrow{\mathbf{O}_1\mathbf{L}}| = \frac{w}{2}$ and $|\mathrm{k'}\overrightarrow{\mathbf{O}_1\mathbf{N}}| = \frac{h}{2}$ we can write

$$k = \frac{w}{2|\overrightarrow{\mathbf{O}_1\mathbf{L}}|}, \tag{10}$$

and

$$k' = \frac{h}{2|\overrightarrow{\mathbf{O}_1\mathbf{N}}|}, \tag{11}$$

therefore

$$\overrightarrow{\mathbf{O}_1} = \overrightarrow{\mathbf{P}} - \frac{w}{2|\overrightarrow{\mathbf{O}_1\mathbf{L}}|} \times \overrightarrow{\mathbf{O}_1\mathbf{L}} - \frac{h}{2|\overrightarrow{\mathbf{O}_1\mathbf{N}}|} \times \overrightarrow{\mathbf{O}_1\mathbf{N}}. \tag{12}$$

Knowing that $\overrightarrow{\mathbf{QR}}$ is parallel to and in the same direction as the x axis of Πi and as $\overrightarrow{\mathbf{O}_1\mathbf{L}}$ is parallel to and in the same direction is on the x axis of Πi we can say that

$$\overrightarrow{\mathbf{O}_1\mathbf{L}} \parallel \overrightarrow{\mathbf{QR}}. \tag{13}$$

Moreover, Knowing that $\overrightarrow{\mathbf{O}_1\mathbf{N}}$ is perpendicular to $\overrightarrow{\mathbf{O}_1\mathbf{L}}$ and $\overrightarrow{\boldsymbol{\alpha}}$; we can write that

$$\overrightarrow{\mathbf{O}_1\mathbf{N}} \parallel \overrightarrow{\boldsymbol{\alpha}} \times \overrightarrow{\mathbf{O}_1\mathbf{L}}, \tag{14}$$

$$\Rightarrow \overrightarrow{\mathbf{O}_1\mathbf{N}} \parallel \overrightarrow{\boldsymbol{\alpha}} \times \overrightarrow{\mathbf{QR}}, \tag{15}$$

where $\overrightarrow{\boldsymbol{\alpha}}$ can be written as the cross product of $\overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{ST}}$.[‡]

Therefore from (12), (13) and (15), we can summarise that the 3D coordinates of $\overrightarrow{\mathbf{O}_1}$ are

$$\boxed{\overrightarrow{\mathbf{O}_1} = \overrightarrow{\mathbf{P}} - \frac{w}{2} \times \widehat{\mathbf{O}_1\mathbf{L}} - \frac{h}{2} \times \widehat{\mathbf{O}_1\mathbf{N}},} \tag{16}$$

where $\overrightarrow{\mathbf{P}} = \overrightarrow{\mathbf{C}} + \widehat{\boldsymbol{\alpha}} \times f$,
where $\overrightarrow{\mathbf{O}_1\mathbf{L}} \parallel \overrightarrow{\mathbf{QR}}$, represents the direction vector of the x axis,
where $\overrightarrow{\mathbf{O}_1\mathbf{N}} \parallel \overrightarrow{\boldsymbol{\alpha}} \times \overrightarrow{\mathbf{QR}}$, represents the direction vector of the y axis,
and where $\overrightarrow{\boldsymbol{\alpha}}$ can be written as the cross product of $\overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{ST}}$.[‡]

The same process can be repeated to find the position of the origin, $\overrightarrow{\mathbf{O'}_1}$. As a result

$$\boxed{\overrightarrow{\mathbf{O'}_1} = \overrightarrow{\mathbf{P'}} - \frac{w}{2} \times \widehat{\mathbf{O'}_1\mathbf{L'}} - \frac{h}{2} \times \widehat{\mathbf{O'}_1\mathbf{N'}},} \tag{17}$$

where $\overrightarrow{\mathbf{P'}} = \overrightarrow{\mathbf{C'}} + \widehat{\boldsymbol{\alpha'}} \times f$,
where $\overrightarrow{\mathbf{O'}_1\mathbf{L'}} \parallel \overrightarrow{\mathbf{Q'R'}}$, represents the direction vector of the x axis,
where $\overrightarrow{\mathbf{O'}_1\mathbf{N'}} \parallel \overrightarrow{\boldsymbol{\alpha'}} \times \overrightarrow{\mathbf{Q'R'}}$, represents the direction vector of the y axis,
and where $\overrightarrow{\boldsymbol{\alpha'}}$ can be written as the cross product of $\overrightarrow{\mathbf{Q'R'}}$ and $\overrightarrow{\mathbf{S'T'}}$ [§].

[‡]When looking at the vector $\overrightarrow{\boldsymbol{\alpha}}$ from a point of view where $\overrightarrow{\mathbf{C}}$ is to the left and $\overrightarrow{\mathbf{C'}}$ is to the right, $\overrightarrow{\boldsymbol{\alpha}}$ must go forwards.
[§]When looking at the vector $\overrightarrow{\boldsymbol{\alpha'}}$ from a point of view where $\overrightarrow{\mathbf{C}}$ is to the left and $\overrightarrow{\mathbf{C'}}$ is to the right, $\overrightarrow{\boldsymbol{\alpha'}}$ must go forwards.

Step 3: Now that we know the position of i, i', $\overrightarrow{\mathbf{C}}$ and $\overrightarrow{\mathbf{C'}}$ relative to the arbitrary origin $\mathbf{O}_2$ of the 3D scene, we can triangulate to retrieve the 3D coordinates of the object.

Let the position vector $\overrightarrow{\mathbf{x}}_i = \begin{pmatrix} x_{\mathbf{x}_i} \\ y_{\mathbf{x}_i} \end{pmatrix}$ on the 2D image plane Πi and the position vector $\overrightarrow{\mathbf{x}}_{i'} = \begin{pmatrix} x_{\mathbf{x}_{i'}} \\ y_{\mathbf{x}_{i'}} \end{pmatrix}$ on the 2D image plane $\Pi i'$ be corresponding points.

First, we need to convert the coordinates of the point $\overrightarrow{\mathbf{x}}_i$ (2D coordinates in the 2D scene) to $\overrightarrow{\mathbf{X}}_i = \begin{pmatrix} X_{\mathbf{X}_i} \\ Y_{\mathbf{X}_i} \\ Z_{\mathbf{X}_i} \end{pmatrix}$, (3D coordinates in the 3D scene).

We know from step 2 the origin of the x and y axes on the plane Πi as well as the direction vector of the x and y axes on the plane Πi.
As a result, in the 3D scene the coordinates of $\overrightarrow{\mathbf{X}}_i$ are

$$\overrightarrow{\mathbf{X}}_i = \overrightarrow{\mathbf{O}}_1 + k \times \overrightarrow{\mathbf{O}_1\mathbf{L}} + k' \times \overrightarrow{\mathbf{O}_1\mathbf{N}}, \tag{18}$$

where $|\mathrm{k}\times\overrightarrow{\mathbf{O}_1\mathbf{L}}| = x_{\mathbf{x}_i}$ and $|\mathrm{k'}\times\overrightarrow{\mathbf{O}_1\mathbf{N}}| = y_{\mathbf{x}_i}$ and k and k' are scalars,
where $\overrightarrow{\mathbf{O}_1\mathbf{L}} \parallel \overrightarrow{\mathbf{QR}}$, represents the direction vector of the x axis,
where $\overrightarrow{\mathbf{O}_1\mathbf{N}} \parallel \overrightarrow{\boldsymbol{\alpha}} \times \overrightarrow{\mathbf{QR}}$, represents the direction vector of the y axis,
and where $\overrightarrow{\boldsymbol{\alpha}}$ can be written as the cross product of $\overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{ST}}$.¶

As $|\mathrm{k}\times\overrightarrow{\mathbf{O}_1\mathbf{L}}| = x_{\mathbf{x}_i}$ and $|\mathrm{k'}\times\overrightarrow{\mathbf{O}_1\mathbf{N}}| = y_{\mathbf{x}_i}$ we can write

$$k = \frac{x_{\mathbf{x}_i}}{|\overrightarrow{\mathbf{O}_1\mathbf{L}}|}, \tag{19}$$

and

$$k' = \frac{y_{\mathbf{x}_i}}{|\overrightarrow{\mathbf{O}_1\mathbf{N}}|}. \tag{20}$$

Therefore in the 3D scene the coordinates of $\overrightarrow{\mathbf{X}}_i$ are

$$\boxed{\overrightarrow{\mathbf{X}}_i = \overrightarrow{\mathbf{O}}_1 \ + \ x_{\mathbf{x}_i} \times \widehat{\mathbf{O}_1\mathbf{L}} \ + \ y_{\mathbf{x}_i} \times \widehat{\mathbf{O}_1\mathbf{N}},} \tag{21}$$

where $\overrightarrow{\mathbf{O}_1\mathbf{L}} \parallel \overrightarrow{\mathbf{QR}}$, represents the direction vector of the x axis,
where $\overrightarrow{\mathbf{O}_1\mathbf{N}} \parallel \overrightarrow{\boldsymbol{\alpha}} \times \overrightarrow{\mathbf{QR}}$, represents the direction vector of the y axis,
and where $\overrightarrow{\boldsymbol{\alpha}}$ can be written as the cross product of $\overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{ST}}$.¶

The same process can be repeated to find the position of the $\overrightarrow{\mathbf{X}}_{i'}$. Hence,

$$\boxed{\overrightarrow{\mathbf{X}}_{i'} = \overrightarrow{\mathbf{O'}}_1 \ + \ x_{\mathbf{x}_{i'}} \times \widehat{\mathbf{O'}_1\mathbf{L'}} \ + \ y_{\mathbf{x}_{i'}} \times \widehat{\mathbf{O'}_1\mathbf{N'}},} \tag{22}$$

where $\overrightarrow{\mathbf{O'}_1\mathbf{L'}} \parallel \overrightarrow{\mathbf{Q'R'}}$, represents the direction vector of the x axis,
where $\overrightarrow{\mathbf{O'}_1\mathbf{N'}} \parallel \overrightarrow{\boldsymbol{\alpha}} \times \overrightarrow{\mathbf{Q'R'}}$, represents the direction vector of the y axis,
and where $\overrightarrow{\boldsymbol{\alpha'}}$ can be written as the cross product of $\overrightarrow{\mathbf{Q'R'}}$ and $\overrightarrow{\mathbf{S'T'}}$.‖

¶When looking at the vector $\overrightarrow{\boldsymbol{\alpha}}$ from a point of view where $\overrightarrow{\mathbf{C}}$ is to the left and $\overrightarrow{\mathbf{C'}}$ is to the right, $\overrightarrow{\boldsymbol{\alpha}}$ must go forwards.
‖When looking at the vector $\overrightarrow{\boldsymbol{\alpha'}}$ from a point of view where $\overrightarrow{\mathbf{C}}$ is to the left and $\overrightarrow{\mathbf{C'}}$ is to the right, $\overrightarrow{\boldsymbol{\alpha'}}$ must go forwards.

Reminder:

$$\overrightarrow{\mathbf{X}} = \overrightarrow{\mathbf{C}} + k \times \overrightarrow{\mathbf{CX}_i}, \tag{23}$$

similarly

$$\overrightarrow{\mathbf{X}} = \overrightarrow{\mathbf{C'}} + k' \times \overrightarrow{\mathbf{C'X}_{i'}}, \tag{24}$$

where k and k' are the solutions to the equation $\overrightarrow{\mathbf{C}} + k \times \overrightarrow{\mathbf{CX}_i} = \overrightarrow{\mathbf{C'}} + k' \times \overrightarrow{\mathbf{C'X}_{i'}}$.

Therefore, from (21),(22),(23) and (24)

$$\overrightarrow{\mathbf{X}} = \overrightarrow{\mathbf{C}} + k \times \overrightarrow{\mathbf{C}[\overrightarrow{\mathbf{O}}_1 \ + \ x_{\mathbf{x}_i} \times \widehat{\mathbf{O}_1\mathbf{L}} \ + \ y_{\mathbf{x}_i} \times \widehat{\mathbf{O}_1\mathbf{N}}]}, \tag{25}$$

similarly

$$\overrightarrow{\mathbf{X}} = \overrightarrow{\mathbf{C'}} + k' \times \overrightarrow{\mathbf{C'}[\overrightarrow{\mathbf{O'}}_1 \ + \ x_{\mathbf{x}_{i'}} \times \widehat{\mathbf{O'}_1\mathbf{L'}} \ + \ y_{\mathbf{x}_{i'}} \times \widehat{\mathbf{O'}_1\mathbf{N'}}]}, \tag{26}$$

where

- k and k' are the solutions to the equation $\overrightarrow{\mathbf{C}} + k \times \overrightarrow{\mathbf{C}[\overrightarrow{\mathbf{O}}_1 \ + \ x_{\mathbf{x}_i} \times \widehat{\mathbf{O}_1\mathbf{L}} \ + \ y_{\mathbf{x}_i} \times \widehat{\mathbf{O}_1\mathbf{N}}]} = \overrightarrow{\mathbf{C'}} + k' \times \overrightarrow{\mathbf{C'}[\overrightarrow{\mathbf{O'}}_1 \ + \ x_{\mathbf{x}_{i'}} \times \widehat{\mathbf{O'}_1\mathbf{L'}} \ + \ y_{\mathbf{x}_{i'}} \times \widehat{\mathbf{O'}_1\mathbf{N'}}]}$
- $\overrightarrow{\mathbf{O}_1} = (\overrightarrow{\mathbf{C}} + \widehat{\boldsymbol{\alpha}} \times f) - \frac{w}{2} \times \widehat{\mathbf{O}_1\mathbf{L}} - \frac{h}{2} \times \widehat{\mathbf{O}_1\mathbf{N}}$, and $\overrightarrow{\mathbf{O'}_1} = (\overrightarrow{\mathbf{C'}} + \widehat{\boldsymbol{\alpha}'} \times f) - \frac{w}{2} \times \widehat{\mathbf{O'}_1\mathbf{L'}} - \frac{h}{2} \times \widehat{\mathbf{O'}_1\mathbf{N'}}$,
- $\overrightarrow{\mathbf{O}_1\mathbf{L}} \parallel \overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{O'}_1\mathbf{L'}} \parallel \overrightarrow{\mathbf{Q'R'}}$,
- $\overrightarrow{\mathbf{O}_1\mathbf{N}} \parallel \overrightarrow{\boldsymbol{\alpha}} \times \overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{O'}_1\mathbf{N'}} \parallel \overrightarrow{\boldsymbol{\alpha}'} \times \overrightarrow{\mathbf{Q'R'}}$,
- $\overrightarrow{\boldsymbol{\alpha}}$ can be written as the cross product of $\overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{ST}}$.
- $\overrightarrow{\boldsymbol{\alpha}'}$ can be written as the cross product of $\overrightarrow{\mathbf{Q'R'}}$ and $\overrightarrow{\mathbf{S'T'}}$.

Note: When looking at the vector $\overrightarrow{\boldsymbol{\alpha}}$ from a point of view where $\overrightarrow{\mathbf{C}}$ is to the left and $\overrightarrow{\mathbf{C'}}$ is to the right, $\overrightarrow{\boldsymbol{\alpha}}$ must go forwards.
Note 2: When looking at the vector $\overrightarrow{\boldsymbol{\alpha}'}$ from a point of view where $\overrightarrow{\mathbf{C}}$ is to the left and $\overrightarrow{\mathbf{C'}}$ is to the right, $\overrightarrow{\boldsymbol{\alpha}'}$ must go forwards.

This is an important result as using this formulae the coordinates of $\overrightarrow{\mathbf{X}}$ can be derived from the information inputed in the algorithm and the 2D coordinates of the matching points, $\overrightarrow{\mathbf{x}_i}$ and $\overrightarrow{\mathbf{x}_{i'}}$, representing $\overrightarrow{\mathbf{X}}$ on the images i and i'.

4.3 Example: Reconstruction of a Rubik's Cube

To illustrate the theory and formulas developed above we will now go through an example.
We shall try to reconstruct a 3D model of a Rubik's cube from two pictures of this object.

We will derive the first coordinate using the general rquations given in (25) and (26). I will then give the values of the rest of the coordinates to display the result of the reconstruction.

Here is the **input information** to the reconstruction:

- Two images i and i' with a list the coordinates (2D) of their corresponding points. (The origin is at the corner down to the left)

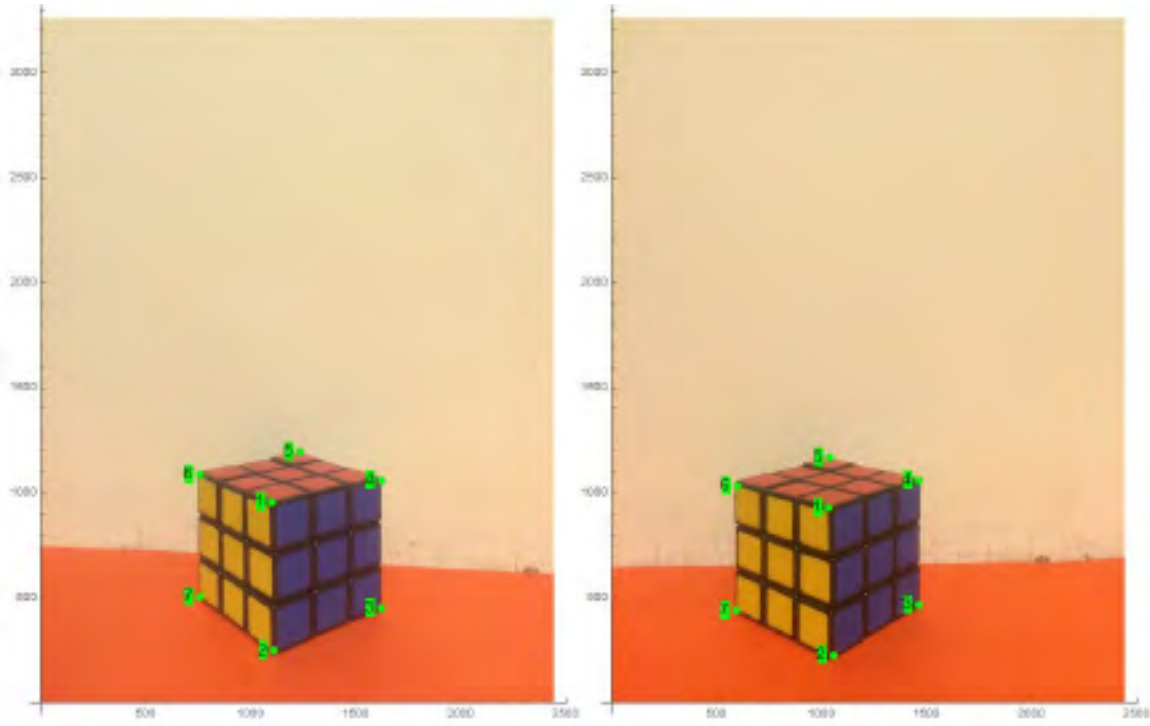


Fig. 4.1. Two images i and i' with manually matched points.
The axes are graduated in pixels (not to scale on those pictures).

Matching Point	**Position Vector on *i* (px)**	**Position Vector on *i'* (px)**
1	$\begin{pmatrix}1097.8\\957.0\end{pmatrix}$	$\begin{pmatrix}1032.6\\929.8\end{pmatrix}$
2	$\begin{pmatrix}1108.7\\250.5\end{pmatrix}$	$\begin{pmatrix}1054.4\\228.7\end{pmatrix}$
3	$\begin{pmatrix}1619.6\\451.5\end{pmatrix}$	$\begin{pmatrix}1462.0\\467.9\end{pmatrix}$
4	$\begin{pmatrix}1619.6\\1060.2\end{pmatrix}$	$\begin{pmatrix}1462.0\\1060.2\end{pmatrix}$
5	$\begin{pmatrix}1233.7\\1196.1\end{pmatrix}$	$\begin{pmatrix}1038.0\\1168.9\end{pmatrix}$
6	$\begin{pmatrix}755.4\\1087.4\end{pmatrix}$	$\begin{pmatrix}597.8\\1033.1\end{pmatrix}$
7	$\begin{pmatrix}755.4\\505.9\end{pmatrix}$	$\begin{pmatrix}592.4\\440.7\end{pmatrix}$

Table. 4.1. Coordinates of the matching points on i and i'.

- The position vectors of the camera centres $\overrightarrow{\mathbf{C}}$ and $\overrightarrow{\mathbf{C'}}$ relative to an arbitrary placed origin, $\overrightarrow{\mathbf{O}_2}$, in a 3D space scene with 3 mutually perpendicular axes: X, Y, Z for i and i' respectively:

$$\overrightarrow{\mathbf{C}} = \begin{pmatrix}0.0\\1.0\\11.5\end{pmatrix},$$

$$\overrightarrow{\mathbf{C'}} = \begin{pmatrix}4.8\\1.3\\11.5\end{pmatrix}.$$

- The direction vectors $\overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{ST}}$ parallel to the plane of image 1, where $\overrightarrow{\mathbf{QR}}$ is parallel to and in the same direction as the x axis of the plane of image 1.

$$\overrightarrow{\mathbf{QR}} = \begin{pmatrix}-4.8\\0.3\\0.0\end{pmatrix},$$

$$\overrightarrow{\mathbf{ST}} = \begin{pmatrix}0.0\\0.0\\1.0\end{pmatrix}.$$

- The direction vectors $\overrightarrow{\mathbf{Q'R'}}$ and $\overrightarrow{\mathbf{S'T'}}$ parallel to the plane of image 1, where $\overrightarrow{\mathbf{Q'R'}}$ is parallel to and in the same direction as the x axis of the plane of image 2.

$$\overrightarrow{\mathbf{Q'R'}} = \begin{pmatrix} -12.4 \\ -1.6 \\ 0.0 \end{pmatrix},$$

$$\overrightarrow{\mathbf{S'T'}} = \begin{pmatrix} 0.0 \\ 0.0 \\ 1.0 \end{pmatrix}.$$

- The focal length of the camera, f, used to take the images:

$$f = 0.412 \text{ cm}$$

- The size of the images, d taken by the camera:

$$d = 3264 \times 2448 \text{ px (where 1 px} = 1.5\ \mu\text{m}^{**}\text{)},$$
$$\Rightarrow (3264 \times 1.5 \times 10^{-4}) \times (2448 \times 1.5 \times 10^{-4}) = 0.4896 \times 0.3672 \text{ cm}$$

Next, lets find the 3D coordinates for matching point 1.

Reminder:

$$\overrightarrow{\mathbf{X}} = \overrightarrow{\mathbf{C}} + k \times \overrightarrow{\mathbf{C}[\overrightarrow{\mathbf{O}}_1 \ + \ x_{\mathbf{x}_i} \times \widehat{\mathbf{O}_1\mathbf{L}} \ + \ y_{\mathbf{x}_i} \times \widehat{\mathbf{O}_1\mathbf{N}}]}, \tag{27}$$

similarly

$$\overrightarrow{\mathbf{X}} = \overrightarrow{\mathbf{C'}} + k' \times \overrightarrow{\mathbf{C'}[\overrightarrow{\mathbf{O'}}_1 \ + \ x_{\mathbf{x}_{i'}} \times \widehat{\mathbf{O'}_1\mathbf{L'}} \ + \ y_{\mathbf{x}_{i'}} \times \widehat{\mathbf{O'}_1\mathbf{N'}}]}, \tag{28}$$

where

- k and k' are the solutions to the equation
$\overrightarrow{\mathbf{C}} + k \times \overrightarrow{\mathbf{C}[\overrightarrow{\mathbf{O}}_1 \ + \ x_{\mathbf{x}_i} \times \widehat{\mathbf{O}_1\mathbf{L}} \ + \ y_{\mathbf{x}_i} \times \widehat{\mathbf{O}_1\mathbf{N}}]} = \overrightarrow{\mathbf{C'}} + k' \times \overrightarrow{\mathbf{C'}[\overrightarrow{\mathbf{O'}}_1 \ + \ x_{\mathbf{x}_{i'}} \times \widehat{\mathbf{O'}_1\mathbf{L'}} \ + \ y_{\mathbf{x}_{i'}} \times \widehat{\mathbf{O'}_1\mathbf{N'}}]}$
- $\overrightarrow{\mathbf{O}_1} = (\overrightarrow{\mathbf{C}} + \widehat{\boldsymbol{\alpha}} \times f) - \frac{w}{2} \times \widehat{\mathbf{O}_1\mathbf{L}} - \frac{h}{2} \times \widehat{\mathbf{O}_1\mathbf{N}}$, and $\overrightarrow{\mathbf{O'}_1} = (\overrightarrow{\mathbf{C'}} + \widehat{\boldsymbol{\alpha}'} \times f) - \frac{w}{2} \times \widehat{\mathbf{O'}_1\mathbf{L'}} - \frac{h}{2} \times \widehat{\mathbf{O'}_1\mathbf{N'}}$,
- $\overrightarrow{\mathbf{O}_1\mathbf{L}} \parallel \overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{O'}_1\mathbf{L'}} \parallel \overrightarrow{\mathbf{Q'R'}}$,
- $\overrightarrow{\mathbf{O}_1\mathbf{N}} \parallel \overrightarrow{\boldsymbol{\alpha}} \times \overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{O'}_1\mathbf{N}} \parallel \overrightarrow{\boldsymbol{\alpha}'} \times \overrightarrow{\mathbf{Q'R'}}$,
- $\overrightarrow{\boldsymbol{\alpha}}$ can be written as the cross product of $\overrightarrow{\mathbf{QR}}$ and $\overrightarrow{\mathbf{ST}}$.
- $\overrightarrow{\boldsymbol{\alpha}'}$ can be written as the cross product of $\overrightarrow{\mathbf{Q'R'}}$ and $\overrightarrow{\mathbf{S'T'}}$.

Note: When looking at the vector $\overrightarrow{\boldsymbol{\alpha}}$ from a point of view where $\overrightarrow{\mathbf{C}}$ is to the left and $\overrightarrow{\mathbf{C'}}$ is to the right, $\overrightarrow{\boldsymbol{\alpha}}$ must go forwards.
Note 2: When looking at the vector $\overrightarrow{\boldsymbol{\alpha}'}$ from a point of view where $\overrightarrow{\mathbf{C}}$ is to the left and $\overrightarrow{\mathbf{C'}}$ is to the right, $\overrightarrow{\boldsymbol{\alpha}'}$ must go forwards.

Step 1: Finding the value of $\overrightarrow{\boldsymbol{\alpha}}$ and $\overrightarrow{\boldsymbol{\alpha}'}$.

$$\overrightarrow{\boldsymbol{\alpha}} = \begin{pmatrix} -4.8 \\ 0.3 \\ 0.0 \end{pmatrix} \times \begin{pmatrix} 0.0 \\ 0.0 \\ 1.0 \end{pmatrix} = \begin{pmatrix} 0.3 \\ 4.8 \\ 0.0 \end{pmatrix}$$

$$\overrightarrow{\boldsymbol{\alpha}}' = \begin{pmatrix} -12.4 \\ -1.6 \\ 0.0 \end{pmatrix} \times \begin{pmatrix} 0.0 \\ 0.0 \\ 1.0 \end{pmatrix} = \begin{pmatrix} -1.6 \\ 12.4 \\ 0.0 \end{pmatrix}$$

**This depends on the camera

Step 2: Finding the value of $\overrightarrow{\mathbf{O_1N}}$ and $\overrightarrow{\mathbf{O'_1N'}}$.

$$\overrightarrow{\mathbf{O_1N}} = \begin{pmatrix}0.3\\4.8\\0.0\end{pmatrix} \times \begin{pmatrix}-4.8\\0.3\\0.0\end{pmatrix} = \begin{pmatrix}0.0\\0.0\\23.13\end{pmatrix} = \begin{pmatrix}0.0\\0.0\\1.0\end{pmatrix}$$

$$\overrightarrow{\mathbf{O'_1N'}} = \begin{pmatrix}-1.6\\12.4\\0.0\end{pmatrix} \times \begin{pmatrix}-12.4\\-1.6\\0.0\end{pmatrix} = \begin{pmatrix}0.0\\0.0\\156.3\end{pmatrix} = \begin{pmatrix}0.0\\0.0\\1.0\end{pmatrix}$$

Step 3: Finding the value of $\overrightarrow{\mathbf{O_1L}}$ and $\overrightarrow{\mathbf{O'_1L'}}$.

$$\overrightarrow{\mathbf{O_1L}} = \begin{pmatrix}-4.8\\0.3\\0.0\end{pmatrix}$$

$$\overrightarrow{\mathbf{O'_1L'}} = \begin{pmatrix}-12.4\\-1.6\\0.0\end{pmatrix}$$

Step 4: Finding the position of $\overrightarrow{\mathbf{O}_1}$ and $\overrightarrow{\mathbf{O'}_1}$.

$$\overrightarrow{\mathbf{O}_1} = \begin{pmatrix}0.0\\1.0\\11.5\end{pmatrix} + 0.412 \times \frac{\begin{pmatrix}0.3\\4.8\\0.0\end{pmatrix}}{\sqrt{(0.3)^2+(4.8)^2}} - \frac{0.4896}{2} \times \frac{\begin{pmatrix}4.8\\0.3\\0.0\end{pmatrix}}{\sqrt{(4.8)^2+(0.3)^2}} - \frac{0.3672}{2} \times \frac{\begin{pmatrix}0.0\\0.0\\1.0\end{pmatrix}}{\sqrt{(1.0)^2}} = \begin{pmatrix}-0.16\\1.4\\11.3\end{pmatrix}$$

$$\overrightarrow{\mathbf{O'}_1} = \begin{pmatrix}4.8\\11.3\\11.5\end{pmatrix} + 0.412 \times \frac{\begin{pmatrix}1.6\\12.4\\0.0\end{pmatrix}}{\sqrt{(1.6)^2+(12.4)^2}} - \frac{0.4896}{2} \times \frac{\begin{pmatrix}12.4\\1.6\\0.0\end{pmatrix}}{\sqrt{(12.4)^2+(1.6)^2}} - \frac{0.3672}{2} \times \frac{\begin{pmatrix}0.0\\0.0\\1.0\end{pmatrix}}{\sqrt{(1.0)^2}} = \begin{pmatrix}4.6\\1.7\\11.3\end{pmatrix}$$

Step 5: Finding the value of k and k'.

$$\overrightarrow{\mathbf{X}} = \begin{pmatrix}0.0\\1.0\\11.5\end{pmatrix} + k \times [(\begin{pmatrix}-0.16\\1.4\\11.3\end{pmatrix} + 1097.8 \times \frac{1.5}{10^4} \times \frac{\begin{pmatrix}4.8\\0.3\\0.0\end{pmatrix}}{\sqrt{(4.8)^2+(0.3)^2}} + 957.0 \times \frac{1.5}{10^4} \times \frac{\begin{pmatrix}0.0\\0.0\\1.0\end{pmatrix}}{\sqrt{(1.0)^2}}) - \begin{pmatrix}0.0\\1.0\\11.5\end{pmatrix}]$$

Similarly,

$$\overrightarrow{\mathbf{X}} = \begin{pmatrix}4.8\\1.3\\11.5\end{pmatrix} + k' \times [(\begin{pmatrix}4.6\\1.7\\11.3\end{pmatrix} + 1032.6 \times \frac{1.5}{10^4} \times \frac{\begin{pmatrix}12.4\\1.6\\0.0\end{pmatrix}}{\sqrt{(12.4)^2+(1.6)^2}} + 929.8 \times \frac{1.5}{10^4} \times \frac{\begin{pmatrix}0.0\\0.0\\1.0\end{pmatrix}}{\sqrt{(1.0)^2}}) - \begin{pmatrix}4.8\\1.3\\11.5\end{pmatrix}].$$

As a result

$$\begin{pmatrix}0.0\\1.0\\11.5\end{pmatrix} + k \times \begin{pmatrix}0.004\\0.410\\-0.056\end{pmatrix} = \begin{pmatrix}4.8\\1.3\\11.5\end{pmatrix} + k' \times \begin{pmatrix}-0.046\\0.420\\-0.060\end{pmatrix}.$$

This sets up the system of an equation with 2 unknowns and 3 equalities:

$0 + 0.004k = 4.8 - 0.046k'$ (e1),
$1 + 0.410k = 1.3 + 0.420k'$ (e2),
$11.5 - 0.056k = 11.5 - 0.060k'$(e3).

From (e1) and (e2) k = 96.7 and k' = 96.0

However, when substituting this result into e3 the equality doesn't hold: 11.5 - 0.056 × 96.7 = 6.08,

11.5 - 0.060 × 96.0 = 5.74,

and 6.08 ≠ 5.74.

This shows that the vectors are skew, this is due to some uncertainties when recording the input of the algorithm and because of rounding numbers; however we can keep those values for k and k': this will not affect the precision of the final result too much.

Step 6: We can now retrieve the overall 3D coordinate.

$$\overrightarrow{\mathbf{X}}_1 = \begin{pmatrix} 0 \\ 1 \\ 11.5 \end{pmatrix} + 96.7 \times \begin{pmatrix} 0.004 \\ 0.410 \\ -0.056 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 40.7 \\ 6.0 \end{pmatrix}$$

Step 7: For even more precision you can find the coordinate $\overrightarrow{\mathbf{X'}}$ which is the average of the possible values of $\overrightarrow{\mathbf{X}}$ (as there will be different values for $\overrightarrow{\mathbf{X}}$ due to the fact that the vectors are skew).

$$\overrightarrow{\mathbf{X}}_{1'} = \begin{pmatrix} 4.8 \\ 1.3 \\ 11.5 \end{pmatrix} + 5.31\times \begin{pmatrix} -0.046 \\ 0.420 \\ -0.060 \end{pmatrix} = \begin{pmatrix} 0.3 \\ 41.6 \\ 5.7 \end{pmatrix}$$

$$\overrightarrow{\mathbf{X'}}_1 = [\begin{pmatrix} 0.4 \\ 40.7 \\ 6.0 \end{pmatrix} + \begin{pmatrix} 0.3 \\ 41.6 \\ 5.7 \end{pmatrix}]/2 = \begin{pmatrix} 0.4 \\ 41.2 \\ 5.9 \end{pmatrix}$$

We now found the first 3D coordinate of our Rubik's Cube here are the other ones:

$$\overrightarrow{\mathbf{X'}}_2 = \begin{pmatrix} -0.07 \\ 29.42 \\ -3.25 \end{pmatrix}$$

$$\overrightarrow{\mathbf{X'}}_3 = \begin{pmatrix} 5.16 \\ 35.40 \\ -3.14 \end{pmatrix}$$

$$\overrightarrow{\mathbf{X'}}_4 = \begin{pmatrix} 4.85 \\ 34.39 \\ 4.78 \end{pmatrix}$$

$$\overrightarrow{\mathbf{X'}}_5 = \begin{pmatrix} -0.09 \\ 38.40 \\ 5.05 \end{pmatrix}$$

$$\overrightarrow{\mathbf{X'}}_6 = \begin{pmatrix} -5.31 \\ 34.17 \\ 4.38 \end{pmatrix}$$

$$\overrightarrow{\mathbf{X'}}_7 = \begin{pmatrix} -5.31 \\ 34.17 \\ -2.62 \end{pmatrix}$$

When those coordinates are plotted this is the result:

Fig.4.2. This is the result of our reconstruction of the Rubik's cube.

However, if you rotate the cube there is a missing coordinate:

Fig.4.3. One coordinate is missing because it is not visible on the input images: more images of the object are needed.

5 Wolfram Demonstration Project

Passionate about this project, I had the opportunity to participate to the Wolfram Mathematica Summer Camp last summer in Boston, where I conceived a Wolfram Demonstration which illustrates extended essay.
As of writing this essay (21/09/14), the demonstration is not available on the Wolfram demonstrations website yet, as this demonstration has been coded using Mathematica 10 which is not supported by their servers yet; however it will soon be published.

Some screen shots of the demonstration (taken from the demonstration running on my personal computer) can be seen below and the code is available in the appendix.

Fig.5.1. This is the welcome screen of the presentation.

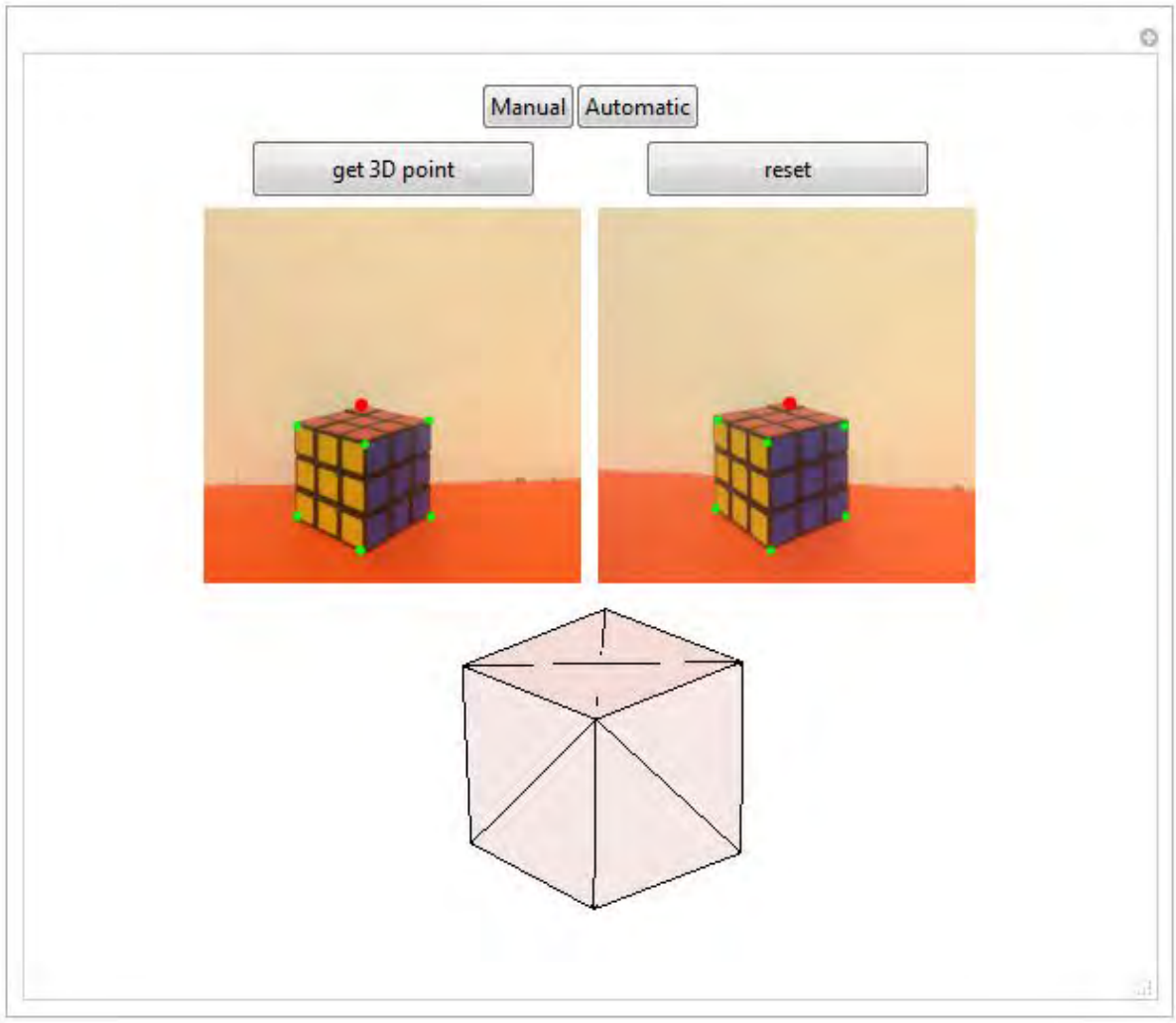


Fig.5.2. This is the reconstruction with a manual selection of the matching points.

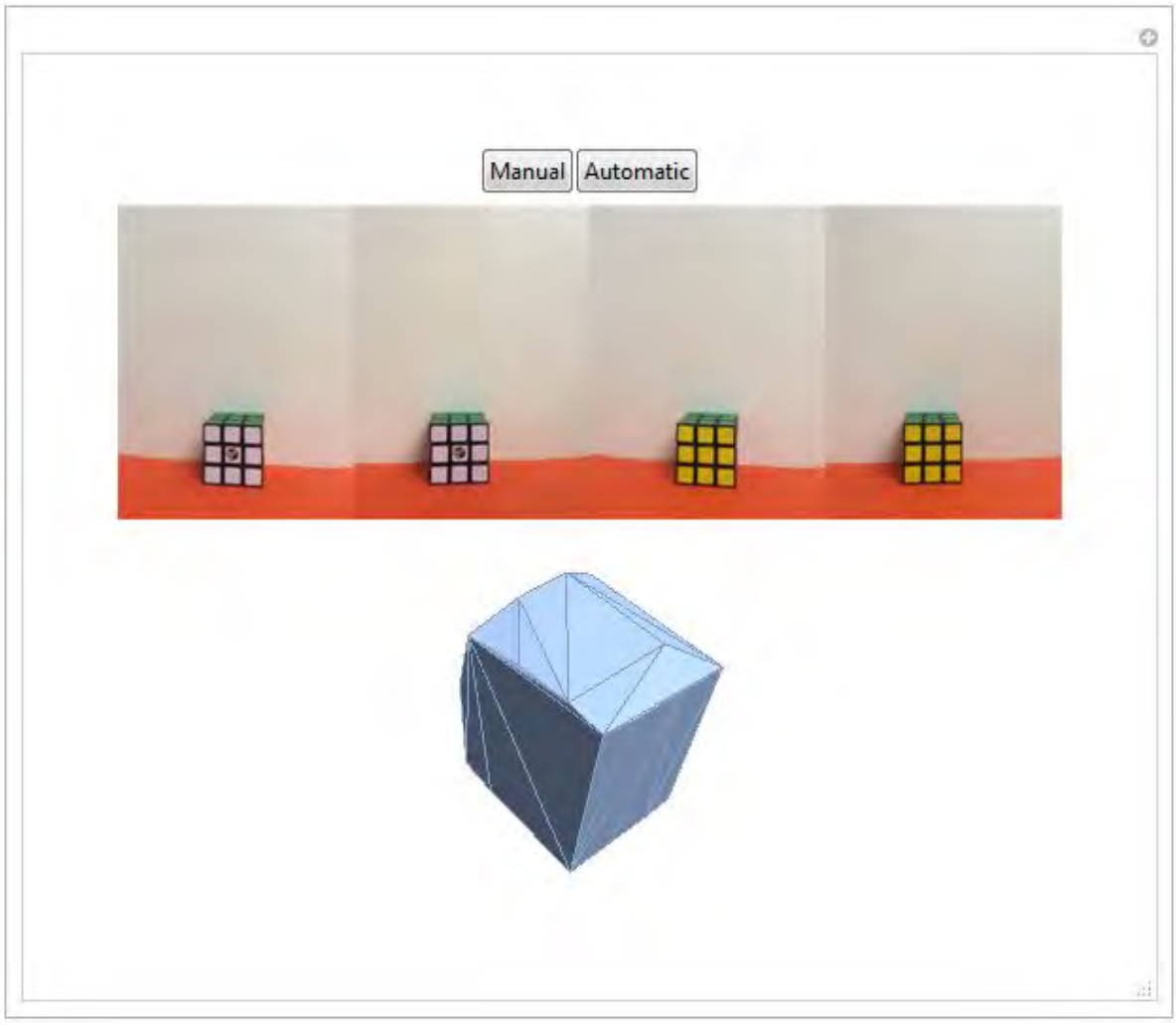


Fig.5.3. This is the reconstruction with an automatic selection of the matching points using the Mathematica function "ImageFeatureTrack".

6 Conclusion

Throughout this extended essay "**A vector approach to 3D reconstruction from multiple views**" has been described, explained and applied using an example and developing a Wolfram Mathematica demonstration project.

Overall this challenge wasn't without any difficulty as when I began researching around the topic, I came across advanced, cutting edge literature which used complex principles such as epipolar geometry and matrices which I had previously never witnessed.
So, I set up a simple base case and gradually extended it to reach the ultimate outcome, using vectors. Achieving to solve this problem was a huge reward and I wish to continue to develop this algorithm in the future.

The initial research question has been resolved however there is still room to improve the algorithm. Indeed, the next steps would be to implement camera auto calibration (no information relative to the position of the camera has to be imputed) and to develop an algorithm to find the corresponding points on the multiple 2D images.
Another expansion to the problem would be to add more views in order to reconstruct the coordinates of the missing corresponding points of the object.

7 Bibliography

References

[1] Cyganek B. and Siebert J. P. *An Introduction to 3D Computer Vision Techniques and Algorithms.* Wiley, 2009.

[2] Hartley R. and Zisserman A. *Multiple View Geometry in Computer Vision.* Cambridge University Press, 2003.

[3] Sunando Sengupta. *Issues in 3D Reconstruction from Multiple Views.* Masters Thesis Indinan Institute of Technology Madras, 2009.

[4] OpenCV-Python Tutorials. *Epipolar Geometry.* `http://opencv-python-tutroals.readthedocs.org/en/latest/py_tutorials/py_calib3d/py_epipolar_geometry/py_epipolar_geometry.html#epipolar-geometry`, last consulted on 21/08/14.

[5] Nikon USA. *Issues in 3D Reconstruction from Multiple Views.* `http://www.nikonusa.com/en/Learn-And-Explore/Article/g3cu6o2o/understanding-focal-length.html`, last consulted on 21/08/14.

[6] Wikipedia. *Pinhole camera.* `http://en.wikipedia.org/wiki/Pinhole_camera`, last consulted on 21/08/14.

8 Appendix

8.1 Code of the Wolfram Mathematica Demonstration

```
cameracentreC[a_, b_, c_] := {a, b, c};
secondPoint[d_, e_, f_] := {d, e, f};
planePiC[a_, b_, c_, d_, e_, f_] := {cameracentreC[a, b, c],
   {secondPoint[d, e, f] - cameracentreC[a, b, c]}, {0, 0, 1}};
vectorA[a_, b_, c_, d_, e_, f_] := {cameracentreC[a, b, c],
   Cross[planePiC[a, b, c, d, e, f][[2]], planePiC[a, b, c, d, e, f][[3]]]};
value1[a_, b_, c_, d_, e_, f_, g_] :=
  Quiet@Solve[x * Norm[vectorA[a, b, c, d, e, f][[2]]] == g, x];
pointC2[a_, b_, c_, d_, e_, f_, g_] := cameracentreC[a, b, c] +
   value1[a, b, c, d, e, f, g][[1, 1, 2]] * vectorA[a, b, c, d, e, f][[2]];
planePiC2[a_, b_, c_, d_, e_, f_, g_] := {pointC2[a, b, c, d, e, f, g],
   {secondPoint[d, e, f] - cameracentreC[a, b, c]}, {0, 0, 1}};
directionofvectorB[a_, b_, c_, d_, e_, f_] :=
  Cross[vectorA[a, b, c, d, e, f][[2]], {0, 0, 1}];
directionofvectorG = Cross[{1, 0, 0}, {0, 1, 0}];
value2[a_, b_, c_, d_, e_, f_, i_] := Quiet@
   Solve[Norm[x * directionofvectorB[a, b, c, d, e, f]] == i / 2 * 1.5 * 10^(-3), x];
value3[h_] := Quiet@Solve[Norm[y * directionofvectorG] == h / 2 * 1.5 * 10^(-3), y];
originoftheImage[a_, b_, c_, d_, e_, f_, g_, h_, i_] := pointC2[a, b, c, d, e, f, g] -
   value2[a, b, c, d, e, f, i][[2, 1, 2]] * directionofvectorB[a, b, c, d, e, f] -
   value3[h][[2, 1, 2]] * directionofvectorG;
printedimageSize = 200;
transformationofCoordinates[i_, j_] := i * 1.5 * 10^(-3) * j / printedimageSize;
valueX[a_, b_, c_, d_, e_, f_, i_, j_] :=
  Quiet@Solve[transformationofCoordinates[i, j][[1]] ==
    Norm[x * directionofvectorB[a, b, c, d, e, f]], x];
valueY[i_, j_] := Quiet@Solve[transformationofCoordinates[i, j][[2]] ==
    Norm[y * directionofvectorG], y];
coordinatesofoverallSystem[a_, b_, c_, d_, e_, f_, g_, h_, i_, j_] :=
  originoftheImage[a, b, c, d, e, f, g, h, i] +
   valueX[a, b, c, d, e, f, i, j][[2, 1, 2]] * directionofvectorB[a, b, c, d, e, f] +
   valueY[i, j][[2, 1, 2]] * directionofvectorG;
directionofRay[a_, b_, c_, d_, e_, f_, g_, h_, i_, j_] :=
  coordinatesofoverallSystem[a, b, c, d, e, f, g, h, i, j] - cameracentreC[a, b, c];
directionofRay[k_, l_, m_, n_, o_, p_, g_, h_, i_, q_] :=
  coordinatesofoverallSystem[k, l, m, n, o, p, g, h, i, q] - cameracentreC[k, l, m];
value4[a_, b_, c_, d_, e_, f_, g_, h_, i_, j_, k_, l_, m_, n_, o_, p_, q_] :=
  Quiet@Solve[{cameracentreC[a, b, c][[1]] +
       x * directionofRay[a, b, c, d, e, f, g, h, i, j][[1]] == cameracentreC[
         k, l, m][[1]] + y * directionofRay[k, l, m, n, o, p, g, h, i, q][[1]],
      cameracentreC[a, b, c][[2]] + x * directionofRay[a, b, c, d, e, f, g, h, i, j][[
         2]] == cameracentreC[k, l, m][[2]] +
       y * directionofRay[k, l, m, n, o, p, g, h, i, q][[2]]}, {x, y}];
overall3DCoordinates[a_, b_, c_, d_, e_, f_, g_, h_, i_, j_, k_,
   l_, m_, n_, o_, p_, q_] := cameracentreC[a, b, c] +
   value4[a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q][[1, 1, 2]] *
    directionofRay[a, b, c, d, e, f, g, h, i, j];
f1[r_, s_] := {ImageResize[r, 200], ImageResize[s, 200]};
f2[r_, s_] := ImageFeatureTrack[f1[r, s]];
f3[a_, b_, c_, d_, e_, f_, g_, h_, i_, k_, l_, m_, n_, o_, p_, r_, s_] :=
  Table[{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p}, {j, f2[r, s][[1]]}];
f4[a_, b_, c_, d_, e_, f_, g_, h_, i_, k_, l_, m_, n_, o_, p_, r_, s_] :=
```

```
  Table[Flatten[{f3[a, b, c, d, e, f, g, h, i, k, l, m, n, o, p, r, s][[x]],
     {f2[r, s][[2, x]]}}, 1], {x, Length[f2[r, s][[2]]]}];
f5[{a_, b_, c_, d_, e_, f_, g_, h_, i_, j_, k_, l_, m_, n_, o_, p_, q_}] :=
  overall3DCoordinates[a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q];
f6[a_, b_, c_, d_, e_, f_, g_, h_, i_, k_, l_, m_, n_, o_, p_, r_, s_] :=
  Table[f5[w], {w, f4[a, b, c, d, e, f, g, h, i, k, l, m, n, o, p, r, s]}];
f7[a1_, b1_, c1_, d1_, e1_, f1_, g1_, h1_, i1_, k1_, l1_, m1_, n1_,
   o1_, p1_, r1_, s1_, a2_, b2_, c2_, d2_, e2_, f2_, g2_, h2_, i2_,
   k2_, l2_, m2_, n2_, o2_, p2_, r2_, s2_] := ConvexHullMesh[
   Join[f6[a1, b1, c1, d1, e1, f1, g1, h1, i1, k1, l1, m1, n1, o1, p1, r1, s1],
    f6[a2, b2, c2, d2, e2, f2, g2, h2, i2, k2, l2, m2, n2, o2, p2, r2, s2]]];
image1[r_] := ImageResize[r, printedimageSize];
j = {{0, 0}};
q = {{0, 0}};
welcomeText = Grid[{{TextCell[Style[
       "Welcome to the 3D Reconstruction from Multiple Views demonstration.
Here are the different modes available:
", Darker[Red], FontFamily → "Helvetica", FontSize → 18],
      TextAlignment → Center], SpanFromLeft}, {TextCell[Style["Manual
", FontFamily → "Helvetica", FontSize → 14, Darker[Gray]],
      TextAlignment → Center], TextCell[Style["Automatic
", FontFamily → "Helvetica", Darker[Gray], FontSize → 14]]},
    {TextCell["You select the corresponding points between the images
        and the demonstration reconstructs the 3D object.

1: Select a pair of corresponding points on the
        images by clicking on them (a red spot will appear)

2: Click on the button (Get 3D coordinates)

3: The 3D point will be constructed (3
        points minimum required to see the reconstruction)
4: Repeat the process for as many corresponding points as you wish.",
      FontFamily → "Helvetica", FontSize → 12],
     TextCell["The computer finds the corresponding points
        for you (don't need to do anything)",
      FontFamily → "Helvetica", FontSize → 12]}}, Alignment → Center];
Auto = f7[28.34, 3.33, 11.5, 29.93, -7.62, 11.5, 4.12, 3264, 2448,

   29.14, -1.27, 11.5, 28.66, -7.62, 11.5,  ,  ,

   -27.34, -3.33, 11.5, -28.93, 7.62, 11.5, 4.12, 3264, 2448,

   -28.14, 1.27, 11.5, -27.66, 7.62, 11.5,  ,  ];
mymanipulate[a_, b_, c_, d_, e_, f_, g_, h_, i_, k_, l_,
   m_, n_, o_, p_, r_, s_, r1_, s1_, r2_, s2_] :=
  Column[{Row[{Button["Manual",
```

```
(image = Grid[{
    {Button["get 3D point", AppendTo[coordinates, overall3DCoordinates[a,
        b, c, d, e, f, g, h, i, j[[1]], k, l, m, n, o, p, q[[1]]]],
      ImageSize → {150, 30}], Button["reset", pts2DL = pts2DR = {};
      coordinates = {}, ImageSize → {150, 30}]},
    {
     ClickPane[Dynamic[Show[{
         ImageCrop[image1[r], {200, 200}, Top],
         Graphics[{PointSize[Medium], Green,
           Point /@ pts2DL, Red, PointSize[Large], Point[#] & /@ j}]
        }]], (AppendTo[pts2DL, #]; j = {#}) &],
     ClickPane[Dynamic[Show[{
         ImageCrop[image1[s], {200, 200}, Top],
         Graphics[{PointSize[Medium], Green,
           Point /@ pts2DR, Red, PointSize[Large], Point[#] & /@ q}]
        }]], (AppendTo[pts2DR, #]; q = {#}) &]},
     {Dynamic[Graphics3D[
        Polygon[Permutations[coordinates]], ImageSize → {200, 200},
        PlotRange → All, Boxed -> False]], SpanFromLeft}
    }]],
  Button["Automatic", (image = Column[{Show[
      ImageAssemble[{image1[r1], image1[s1], image1[r2], image1[s2]}],
      ImageSize → 500], Show[Auto, ImageSize → {200, 200}]},
    Alignment → Center])]], Dynamic[image]}, Alignment → Center];

Manipulate[mymanipulate[0, 1.27 - 0.25, 11.5, -4.752,
  1.587 - 0.25, 11.5, 4.12, 3264, 2448, 4.752, 1.5875 - 0.25, 11.5, -7.62,
  -0.25, 11.5,  ],
 {j, ControlType → None}, {q, ControlType → None},
 Initialization ⇒ {pts2DL = pts2DR = {}, coordinates = {}, image = welcomeText},
 ContentSize → {600, 500}, Alignment → Center, SaveDefinitions → True]
```

What are the factors influencing HIV prevalence in Tanzania?

The IB introduced a new form of the Extended Essay in 2013 called the World Studies Extended Essay. Interdisciplinary in nature, the WSEE expects students to start from some issue of global significance (in this case the transmission of HIV/AIDS) and explore it in a local context through the lens of two IB subjects. The goal of the essay then is to highlight the need of rigorous academic investigation of important social issues, the inherent interdisciplinary of genuine solutions to these real-world problems and the importance of bringing to life the student's own classroom learning.

Emma Harrington - World Studies

I am very glad I decided to write a World Studies essay and it was a genuinely rewarding experience. The opportunity to look at a virus from the perspective of humanities raised a lot of questions for me. It gave me an insight to the impact that social factors have on the nature of disease and the importance of culture and education in the health of a population. I was able to appreciate both the scale of the HIV/AIDs problem on a national and global scale, but also the significance it has to an individual. Data was challenging to come across, largely due to the stigma surrounding HIV and the challenges of studying a developing country on a relatively small scale, so starting was the hardest part. Writing my Extended Essay gave me a chance to really explore my own interest for this area and after a long time spent researching, writing and cutting 6000 words to 4000, I was pleased with the outcome.

Supervisor: Alex Patton

As a future medic Emma's EE was motivated by an interest in disease, but also by a family trip to Tanzania, which led to her investigating the Epidemiology of HIV in the country. This fitted the brief for a World Studies Essay in that it required the analytical perspectives of two disciplines – Biology and Geography. However, a paucity of publically available data which that was reliable, accurate and at the correct spatial scale made the reality of project very challenging indeed. Emma overcome this through dogged determination and the result is an interesting investigation that if taken further could have important consequences for public policy in the country.

Contents

Abstract

HIV plays some significance in all nations. Whilst the epidemic is under control in many wealthy countries, it remains a serious issue in most of sub-Saharan Africa. Understanding the virus and the factors that shape the epidemic is crucial in controlling it. With 5% of Tanzania's population HIV positive on average, yet a variation of 10.5% between the highest and lowest regions, examining the factors that influence the disparity in HIV prevalence in Tanzania is useful in developing this understanding.

The factors influencing HIV prevalence in Tanzania's mainland are primarily biological factors- frequency of intercourse, gender, effectiveness of treatment and the stage of infection of an individual. However, between regions, socioeconomic and geographical factors influence these biological factors and shape the epidemic. Wealth, education and geography seem to be the most important factors in shaping the epidemic. These main factors further influence the socioeconomic characteristics of a society, leading to a variation in HIV prevalence between regions. The importance of each factor varies between regions and relationships between factors influence the way one factor can shape the epidemic. Perhaps surprisingly, HIV prevalence is highest among the wealthiest even though the wealthiest are most educated. However, in less wealthy but educated populations, HIV prevalence is lower.

HIV prevalence is influenced by wealth due to the large disparity in lifestyle between the rich and the poor. The socioeconomic factors influencing HIV prevalence vary between the rich and the poor and so educational programmes do not cover all issues. This disparity means that neither education nor treatment is universal; the epidemic is thus difficult to control.

List of Illustrations and Tables

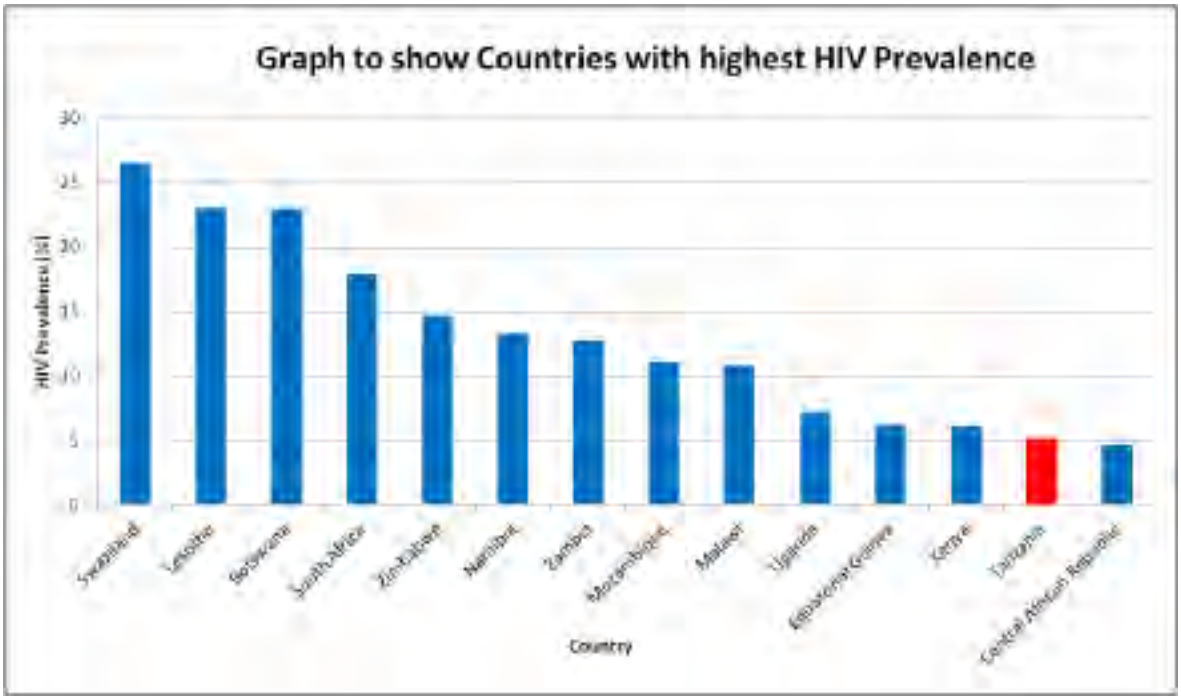
Appendix III: Table to show Spearman's Rank Test of Percentage of Population with Primary Education (minus extremes) vs. HIV Prevalence

Appendix III: Table to show Spearman's Rank Test of Percentage of Population with Primary Education vs. HIV Prevalence

Introduction

Human Immunodeficiency Virus (HIV) infects 33.4 million people worldwide (AIDS, 2012); however there is considerable variation in its impact across the globe. Whilst sub-Saharan Africa accounts for only 13% of the world's population, 64% of all HIV/AIDS sufferers live within this region (Whiteside, 2008, p.6). Tanzania ranks 13th in the world by HIV prevalence (Figure 1) (proportion of the population HIV positive), with 5% of the population aged 12-49 infected (CIA, 2012). In 1983, the first cases of HIV/AIDS were reported in Tanzania's Kagera region. By 1987 the virus had spread to all administrative regions (AVERT, 2012). Each region of Tanzania is impacted by HIV; however the extent of this varies considerably with prevalence ranging from 2.9% in some to 12% in others. Understanding the factors causing this variation is key in developing a strategy to control the epidemic. Fundamentally, the factors influencing HIV prevalence are biological factors which increase the likelihood of transmission; the stage of infection, frequency of intercourse, gender and use of treatment.

Figure 1: Graph to show countries with highest HIV Prevalence



(Emma Harrington, 2014)

Pathology

HIV is a retrovirus and belongs to a family of viruses known as lentiviruses which are characterised by the slow progressive nature of disease once an individual has been infected (Whiteside, 2008, p.22). The virus is present in all bodily secretions, particularly the blood, semen and cervical and vaginal secretions. These secretions are the carriers for HIV in the three main ways of transmission. HIV is most commonly transmitted through unprotected intercourse with an infected individual (accounts for 90% of transmission globally), from an infected mother to her unborn child and by contaminated blood and blood products (Greenwood et al, 2007, p.561).

Once transmitted, HIV targets CD4 (Helper T) cells and becomes established in these cells by integrating its own genetic material into the DNA of the cell. HIV consists of 2 strands of viral RNA which, on entering the host cell, are copied into DNA by the enzyme reverse transcriptase. Integrase then inserts viral DNA into the DNA of the host cell. When this cell is activated, it will produce virus proteins which then form a new copy of the HIV. If this process happens very rapidly and in high numbers, the host cell may die by rupturing. The cell may also perish by cytotoxic T cell lysis[1] of infected cells, attack of the infected helper T cell by natural killer cells (Greenwood et al, 2007, p.557), as well as destruction by HIV specific CD8 cytotoxic lymphocytes (Tan, 2001, p.690).

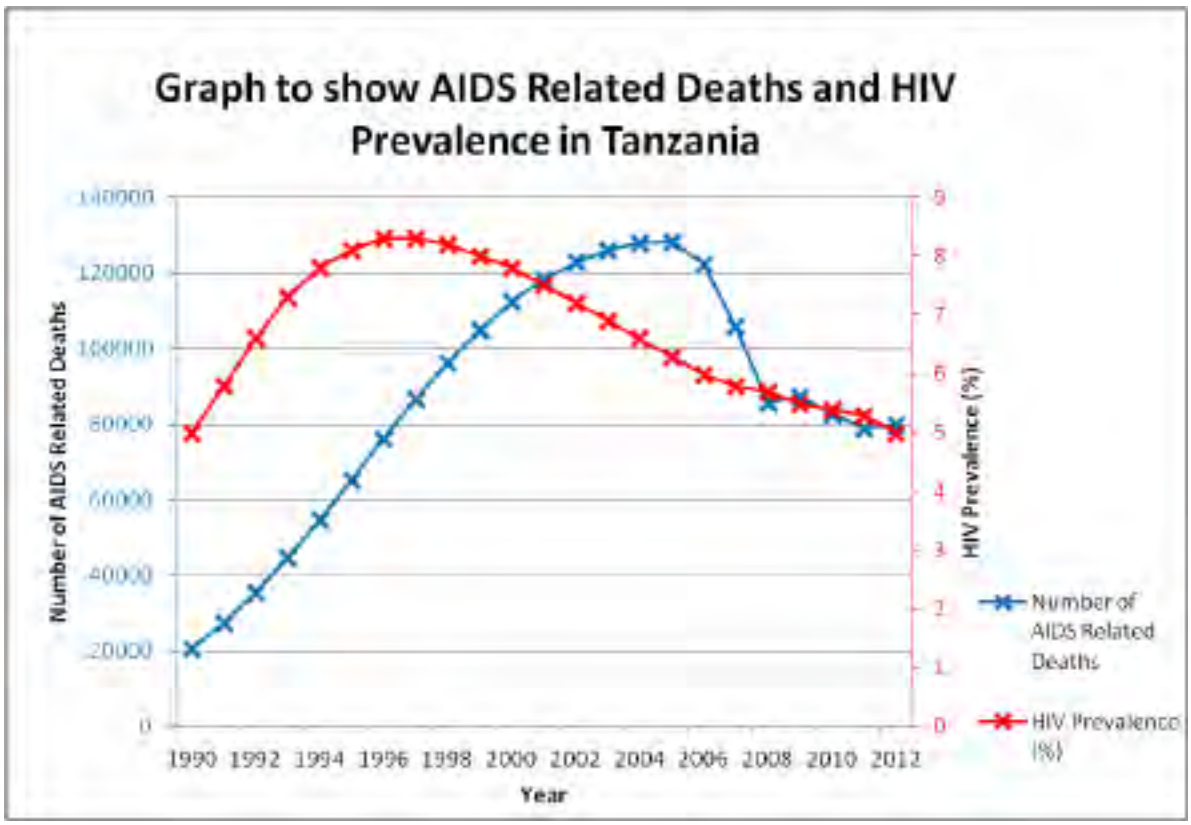
CD4 cells have 2 main functions; to stimulate other leucocytes to attack pathogens and promote an inflammatory response and to release lymphokines to trigger growth and differentiation of B and T cells. Depletion of CD4 cells reduces the body's ability to fight both HIV and other pathogens. This leaves the body susceptible to disease, particularly a group of diseases; collectively forming the syndrome, AIDS. Inability to fight these diseases is the cause of death in an AIDS patient.

During the first 6-9 weeks of infection the virus replicates rapidly; the viral load peaks and the CD4 cell count begins to decrease (Whiteside, 2008, p.27). The body begins to coordinate an immune response to the virus; CD4 cells are produced rapidly however, are destroyed at an equal rate. This period is referred to as the incubation stage and typically lasts 10 years (AIDS, 2013). During this stage, the HIV count is almost undetectable and remains constant. Throughout the incubation stage, the CD4 count does gradually fall until it is so low that the body can no longer

coordinate an effective immune response and the HIV count rapidly rises, leading to death.

Since 1990 the HIV prevalence in Tanzania rose to 8.3% in 1996/97 and has subsequently fallen to 5% (The World Bank, 2012). The prevalence peaked in 1992 and deaths by AIDS 13 years later (Figure 2). This delay represents the incubation stage of infection and the delayed onset of AIDS.

Figure 2: Graph to show AIDS related Deaths and HIV Prevalence in Tanzania

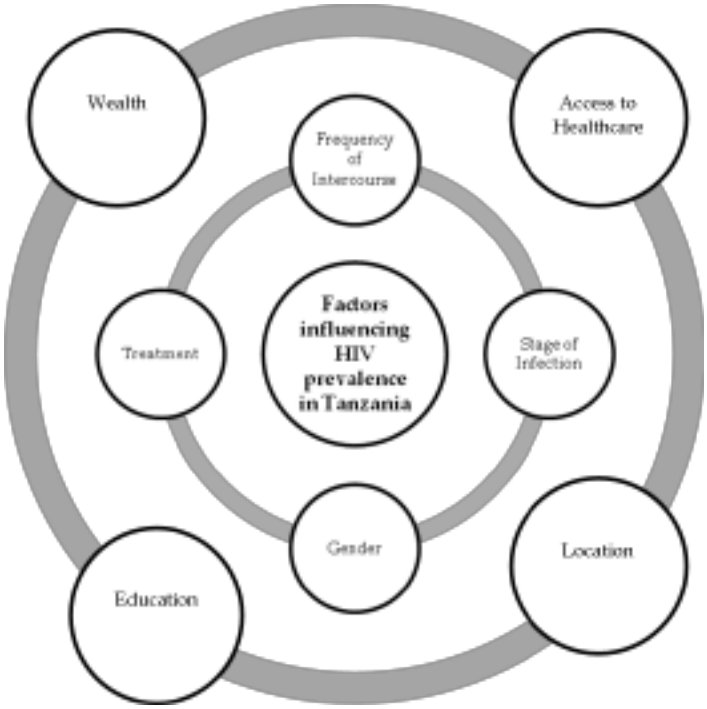


(Emma Harrington, 2014)

What are the factors influencing HIV/AIDS prevalence in Tanzania?

Whilst biological factors are the initial determinants in an HIV epidemic, socioeconomic and geographical factors influence these, shaping the epidemic. These factors are all related and influence each other. Combined, they result in the nature of the HIV epidemic in Tanzania (Figure 3).

Figure 3: The factors influencing HIV prevalence in Tanzania



(Emma Harrington, 2014)

Methodology

After spending some time in an HIV clinic it became apparent that many HIV patients originate from Tanzania and the rest of East Africa. Following a literature search, I identified some Tanzania HIV data. This, combined with my own interest in the country after visiting as a tourist, led to my wish to understand the HIV epidemic in Tanzania.

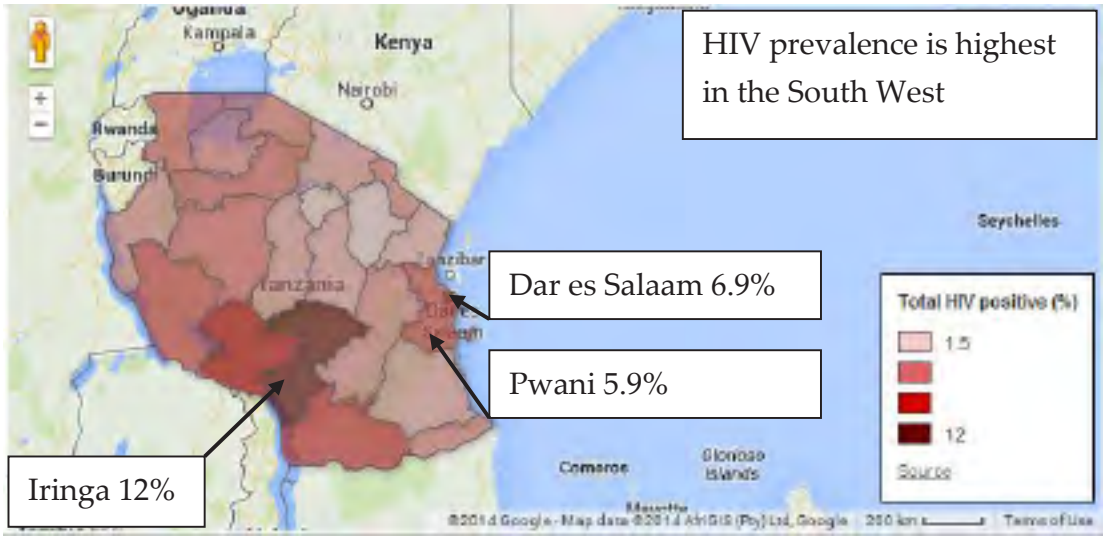
The majority of data used to analyse the trends in prevalence, and factors influencing it, was extracted from the 2011-2012 HIV/AIDS and Malaria Indicator Survey. Data was also sourced from the World Health Organisation and the Tanzania Poverty and Human Development Report. The data was transferred into a spreadsheet and adjusted to fit the Tanzanian regions prior to the 2012 changes. The data was also altered to suit the factors, for example data for average educational attainment for the entire population was not available and values for males and females were averaged to overcome this issue. The data was transferred onto maps, created using Google Fusion Tables, to visualise trends.

Variation in HIV prevalence

Figure 4: Administrative regions of Tanzania

(Wikipedia, 2011)

Figure 5: Total HIV Prevalence

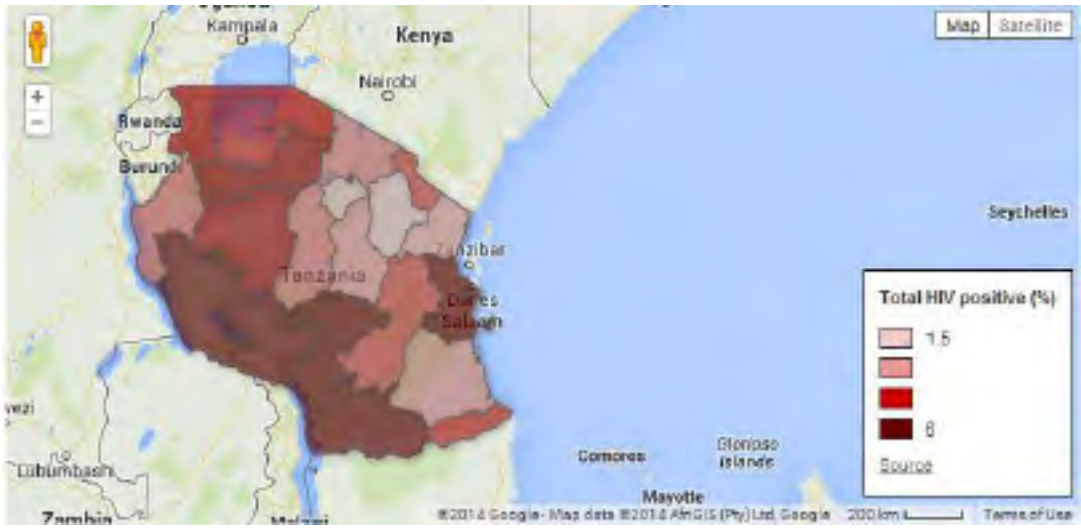


(Emma Harrington, 2014)

If the upper extremes are removed from the map, a geographical divide in HIV prevalence between the East and West becomes more apparent (Figure 6, Figure 7). Prevalence is highest in the South West, followed by the North West and then the eastern side of the country, in particular the North East. The Manyara region is least

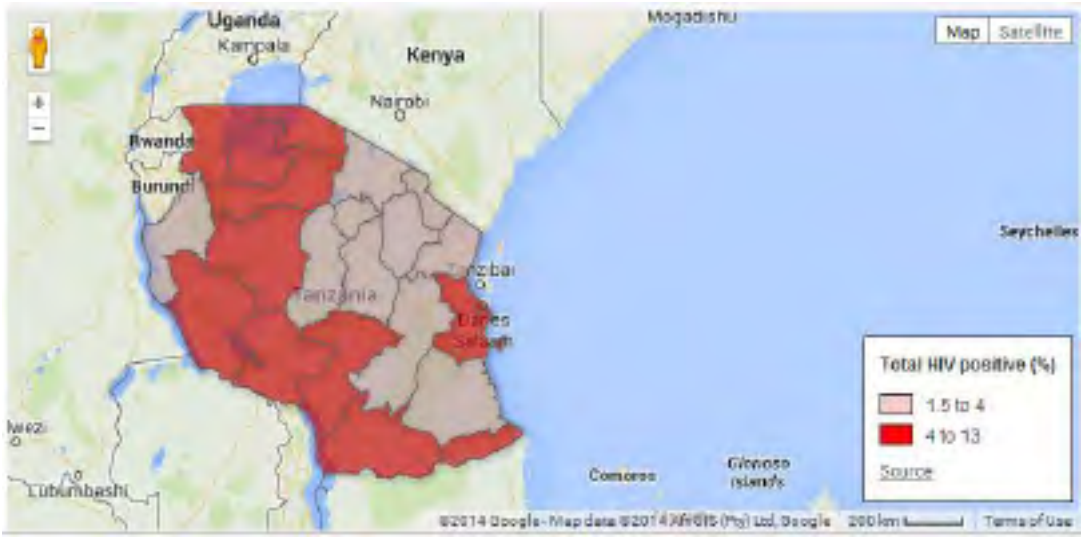
affected by HIV with only 1.5% of the population tested positive. Prevalence is also low in Dodoma and Tanga.

Figure 6: HIV Prevalence (minus extremes, range: 1.5%-6.0%)



(Emma Harrington, 2014)

Figure 7: Total HIV Prevalence (bucket shading)



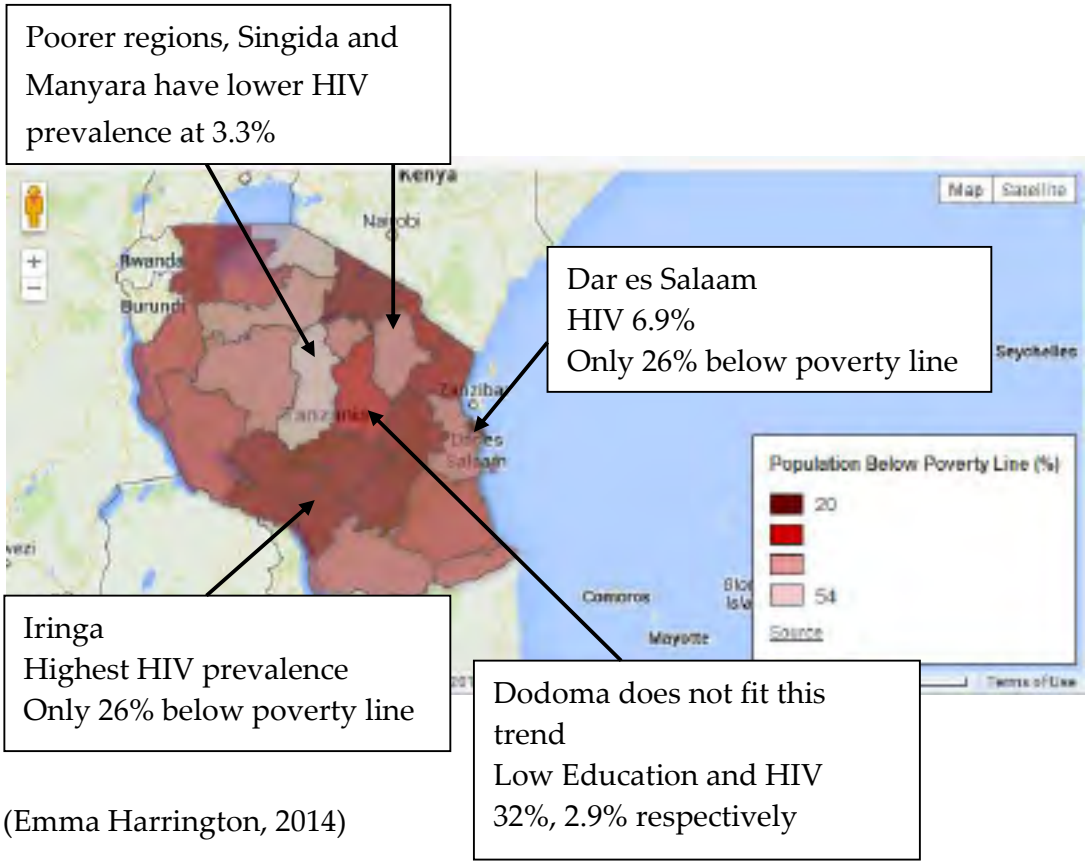
(Emma Harrington, 2014)

Wealth

One of the key factors driving variation in HIV prevalence is wealth. Nationally, the HIV epidemic in Tanzania has been established by poverty and a lack of response to the virus (Whiteside, 2008, p.40) (Tanzania Commission for HIV/AIDS, 2003).

However, within the country, HIV prevalence is exacerbated by wealth. HIV prevalence among the highest wealth quintile is 6.6%, 2.6 higher the lowest. This is also apparent on a regional scale (Figure 8).

Figure 8: Proportion of Population below the Poverty Line



(Emma Harrington, 2014)

Mobility and Interactions

Wealth relates to urban residence and HIV prevalence is higher in urban than rural areas (7.2% compared to 4.3%). In Tanzania, the majority of urban residents are in the nation's highest wealth quintile, whilst in rural areas almost 50% of the population are in the lower 2. With larger social networks and increased interactions between people in cities, HIV prevalence is inevitably higher. HIV prevalence in urban regions is higher such as 6.9% in Dar es Salaam, one of Tanzania's largest cities. Wealthy individuals, rural and urban, are likely to have increased mobility allowing the development of social networks. Wealth is even more crucial here as it not only allows access to more partners, but can reduce the time between partners by access to transport and money (Shelton et al, 2005 p. 1058). This is key as in the first stages of infection the viral load is very high and transmission is much more likely, so the virus can begin to spread within a growing social network, quickly building momentum and increasing HIV prevalence.

Access to Treatment

There is currently no cure for HIV and the most widely available treatment is Anti Retroviral Therapy (ART). There are 5 classes of drug available used in combination which prevent the virus from replicating (Welch, 2014). ART is expensive and are not suited to all patients. Commencing treatment may be very time consuming as patients are often required to try a number of drug combinations. A well established healthcare system, which patients can access on a regular basis, is thus key in treating the virus. According to the WHO, in 2011, only 40% of Tanzanian's requiring ART were receiving treatment (AVERT, 2012) and it's likely that those few receiving it are the wealthier people who have the time and resources available to access clinics. Most Tanzanians access healthcare facilities on foot. Over 95% of urban residents do not need to travel more than 5km to a facility and the majority travel less than 2km. Conversely, in poorer rural areas the majority travel 2-10km. Patients, especially those with symptoms, are less likely to make this journey. ART increases the length of the incubation stage of infection and delays the onset of AIDS. In a population where most infected individuals can access this treatment, there will be more people surviving HIV and thus prevalence will increase. It is important to note that ART decreases the viral count and thus decreases likelihood of transmission but does not eliminate it completely. In a population where few people have treatment, they will die early from the disease and will reduce the prevalence. This is a key reason why wealthier populations in Tanzania show a higher HIV prevalence. Moreover, access to ART will reduce any symptoms of the disease leaving the individual sexually active and much more likely to pass on the virus than an individual who has not received treatment.

Education

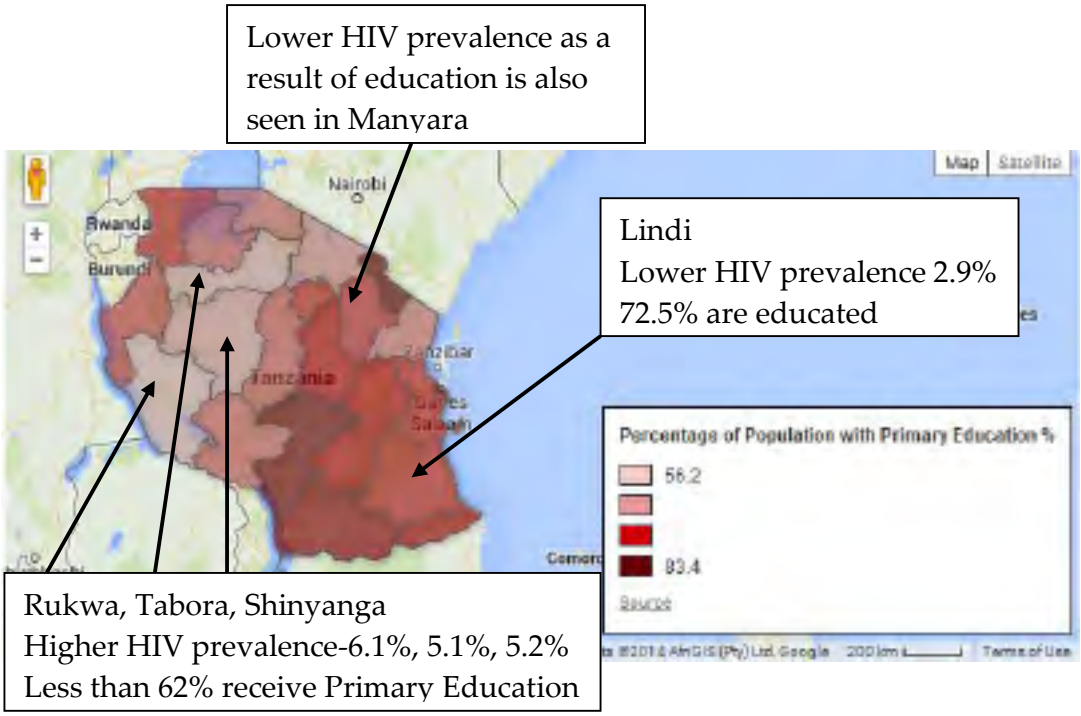
Whilst wealth is central in influencing levels of HIV, education plays a key role. Educational attainment varies considerably within Tanzania, ranging from 56.2% of the Rukwa population obtaining primary education to 83.5% in Dar es Salaam. Eastern Tanzania tends to be more educated than the west however; there are a few anomalies to this. 72.1% of the Kagera population, situated in the North West, is educated to primary level; this figure is relatively high compared to surrounding regions (Figure 9).

Some regions in which a high proportion of the population are educated, such as Dar es Salaam and Iringa, also experience high levels of HIV prevalence. This is most probably attributable to wealth as, although the wealthiest have the widest access to

education, HIV prevalence is high for the socioeconomic factors previously stated. Furthermore, higher levels of wealth in these regions may be as a result of wide access to education.

Aside from the wealthiest regions (Dar es Salaam, Ruvuma, Iringa, Arusha), areas where education levels are higher have a lower HIV prevalence (figure 9). Once these anomalies are excluded a spearman's rank test shows a weak negative correlation (see appendix III).

Figure 9: Proportion of Population with Primary Education



(Emma Harrington, 2014)

Government Attention and Health Education

The relationship between education and HIV is to be expected. There is significant government attention towards raising awareness of HIV and magazines such as Fema and Si Mchezo are provided in schools, discussing issues such as HIV prevention (AVERT, 2012). Schooling is a clear source of HIV knowledge; 20% of the uneducated population have comprehensive knowledge[2] of the virus compared to

[2] 'Comprehensive knowledge' is knowing that consistent use of condoms during sexual intercourse and having one uninfected faithful partner can reduce the chance of infection, knowing that a healthy looking person can have HIV and rejecting the two most common

double this for those with primary education. Almost 100% of the population, with or without an education, are aware of the virus however, a school education is important in establishing more detailed knowledge of HIV and its prevention. 73% of men and women know that risk of HIV infection can be reduced by using a condom but only 57% of those aged 15-24 with no education know where to source condoms compared to 76% of those with primary education and 82% with secondary education.

Stigma

Education influences attitudes towards HIV. A widespread issue in tackling the HIV epidemic is the stigma associated with the virus and the fear of discrimination experienced by those infected. This can result in a reluctance to disclose HIV status which delays treatment and reduces the likelihood of preventative behaviour, collectively increasing the chance of HIV transmission. The most educated tend to be the most accepting when it comes to HIV. Only 13.7% of the uneducated population were accepting in all 4 indicators[3] compared to 33.9% with primary education and 50.9% with secondary education. HIV education in schools reduces the stigma associated with infection and contributes to the reduced HIV prevalence among the more educated regions.

Behaviour

Increased knowledge of HIV also has an influence on the sexual behaviours of society. Polygyny and concurrent relationships are fairly common practice across Tanzania and facilitate the spread of HIV. The proportion of uneducated men with 2+ wives, 16.9%, is more than double that of men with secondary education. Concurrent relations allow the spread of HIV through the development of a large web of sexual relationships (Figure 10). Moreover, short intervals between sexual contact with an infected individual and contact with another partner increases the chance of transmission as the viral load is higher in this early stage of infection.

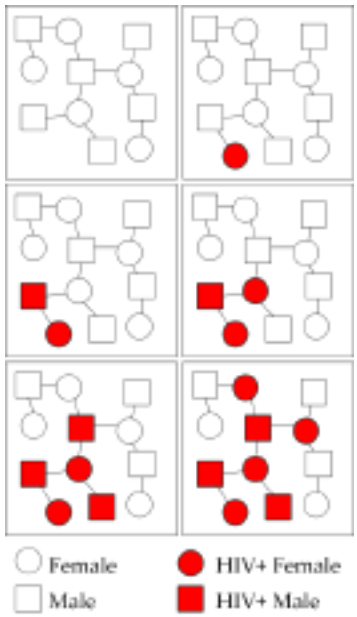
---

local misconceptions about HIV transmission-HIV can be transmitted through mosquito bites and by supernatural means.

[3] Members of the population were asked about 4 indicators and were recorded as expressing an accepting opinion or the contrary. The respondents were asked if they: are willing to take care of a family member with the AIDS virus in the home, would buy fresh vegetables from a shopkeeper with the virus, think a female teacher who is HIV positive but not sick should continue teaching, would not keep it secret if a family member was infected with HIV.

Figure 10: The rapid spread of HIV as a result of concurrent relationships and development of an increasing sexual network

(Emma Harrington, 2014)

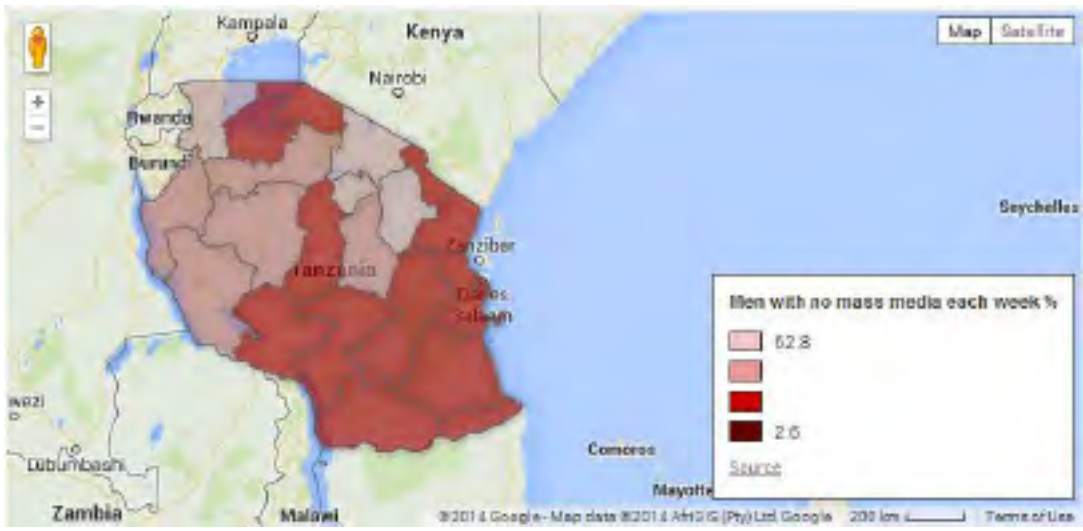


Moreover, if the youth remain in school longer, they are likely to first have intercourse later. In Tanzania, the group with the highest proportion who have intercourse before the age of 15 is those without education. As educational attainment increases, the proportion of the population in this category decreases. If the population remain in school for longer, delaying intercourse, not only do they benefit from the knowledge provided about prevention but are more mature and more likely to have control over the use of contraception, reducing the chances of HIV infection when they do later have intercourse.

Other Forms of Education

Outside of schools, prevention programmes run through newspapers, radio and television. Femina HIP, a government led programme, runs weekly talk shows about HIV (AVERT, 2012). There seems to be a correlation between access to these forms of mass media and HIV prevalence. In general, women have minimal access to mass media however; a trend is seen among men (Figure 11). Again, wealth plays a role here. In wealthy regions such as Iringa, Mbeya and Dar es Salaam, although there is wide access to mass media HIV prevalence is still high. However, in less wealthy regions where there is access to mass media, HIV prevalence tends to be lower, evident in Mtwara, Lindi and Morogoro. In Shinyanga, there is less access to media and as a result HIV prevalence is higher. This trend is not clear in all regions. Arusha, Manyara and Dodoma have low HIV prevalence (3.2%, 1.5%, 2.9%) but also do not have wide access to mass media. These anomalies are perhaps explained by access to school education and a lack of wealth in these areas, leading to a lower HIV prevalence.

Figure 11: Proportion of Men with no access to Mass Media each Week



(Emma Harrington, 2014)

The HIV epidemic in Tanzania is shaped by these socioeconomic factors however is also influenced by the country's geographical characteristics.

Geography

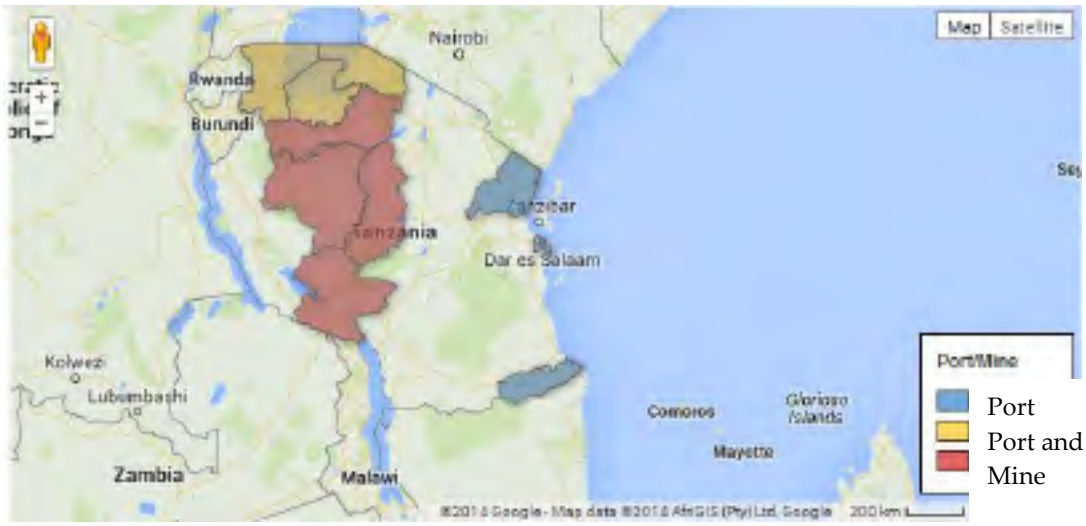
Whilst high HIV prevalence in many regions can be explained by wealth, a number of Northern regions have a higher HIV prevalence than others yet do not appear to be wealthier.

International Trade

Tanzania's large coastline and the presence of Lake Victoria have allowed the development of ports (Figure 12). Kenya, Uganda and North Tanzania border with Lake Victoria. Spanning across 3 Tanzanian regions, Mara, Mwanza and Kagera, Lake Victoria is a key location for trade. HIV prevalence is higher in some of these regions, on average 4.5% in the regions around Lake Victoria. Whilst this value is still below the national average, it is higher than prevalence for much of the East. The presence of a sea port in Tanga does not seem to cause increased HIV prevalence. This is most likely due to the fact that 95% of international sea trade occurs from Dar es Salaam (Tanzania Ports Authority, 2012). In general, these areas involved in international trade have an increased HIV prevalence due to the movement of people. The ports surrounding Lake Victoria have connections to Uganda and Kenya, countries with high HIV prevalence. These connections not only mean that more cases of HIV enter the country but also mean the population has

increased mobility and people are more likely to come in contact with infected individuals, facilitating the spread of HIV.

Figure 12: Distribution of Ports and Mines by region



(Emma Harrington, 2014)

Mining

Mining, dominated by gold extraction, contributes to 4.6% of Tanzania's economy (All Africa, 2012) and is concentrated to the West. Three large mines, Bulyanhulu (Miners Weekly, 2005), Buzwagi and Kirondatal are located in Shinyanga and additional mines are found in the regions indicated (Figure 12). The mining trade is characterised by the movement of male workers away from home to isolated mining sites. This often results in the development of brothels as men move away from their partners. Frequency of paid sex is slightly higher in these regions and the use of a condom is less common during paid intercourse (53%) compared to 60% for any intercourse. In an area where the majority of the population is sexually active and of working age and income is mostly generated by men, male dominance over women may develop. This may lead to the spread of HIV as the use of contraception is minimal and men have shorter intervals between partners. In this case the HIV transmission is driven by increased sexual activity in an area, worsened by a lack of contraception and infection is made more likely as a large part of the population is in the early stages of infection when the viral load is high. The presence of mining in these regions is a possible explanation for the increased HIV prevalence (Shinyanga 5.2%, Mbeya 9.0%).

Other means of Transmission

Whilst 80% of HIV infections in Tanzania are transmitted by heterosexual intercourse (AVERT, 2012), a significant proportion are passed on via two other means of transmission; through the use of contaminated needles by Injecting Drug Users (IDUs) and from mother-to-child.

Injecting Drug Users

There are approximately 25,000 IDUs in Tanzania and 50% are expected to have a positive HIV status (AVERT, 2012). Though there is minimal data concerning drug use, it is probable that IDUs are concentrated to urban areas where access to drugs and a community of drug users is more established. This perhaps contributes to higher HIV prevalence in urban, wealthy regions. There is probably greater access to drugs in regions such as Shinyanga and Dar es Salaam due to the presence of international ports compared to poorer, rural regions where drug use is probably minimal.

Mother-to-child Transmission

Almost 1/5th of HIV infections in Tanzania are due to mother-to-child transmission. This is almost entirely avoidable with ART for mother and child as well as safer feeding practices. (AVERT,2012) Education and access to healthcare are vital in prevention transmissions by this means. Mother-to-child transmission may contribute to higher HIV prevalence in uneducated regions. Although it would be expected that in wealthier regions where access to treatment is more universal, mother-to-child transmission would be minimal and thus HIV prevalence would be lower, this is not the case. This can be explained by the relationship between heterosexual transmission and wealth, increasing prevalence. As the majority of transmission is via heterosexual intercourse, even though mother-to-child transmission is probably lower in wealthy regions, this is not apparent when analysing prevalence data.

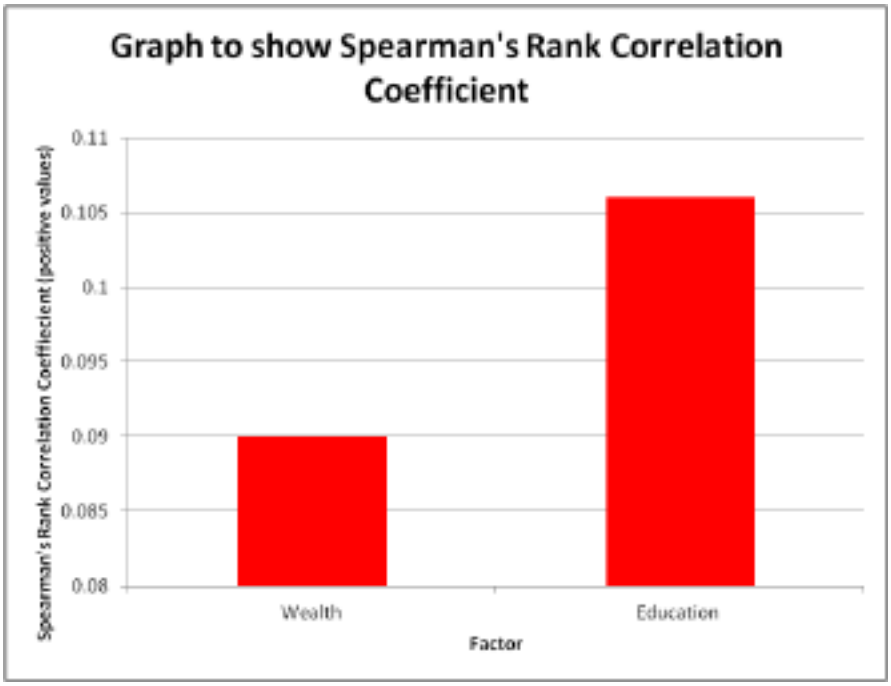
Conclusion and Evaluation

The pathology of HIV, combined with the characteristics of Tanzania's population, has led to the geographical trend in HIV prevalence seen between regions. Levels of wealth and education and the geography of Tanzania shape the epidemic. Each factor tends to shape the epidemic for numerous reasons. Education not only raises awareness of HIV, but influences attitudes towards the epidemic and the sexual behaviour a society. The relative importance of the factors across regions is varied.

Educational attainment is high in Iringa but has little impact on reducing HIV prevalence because high levels of wealth are very much responsible for the increased levels of HIV. Conversely, high levels of education in Morogoro do successfully reduce HIV prevalence. This highlights the complex relationships between the factors controlling the epidemic; the way that the extent to which one factor shapes the epidemic is controlled by the relative presence of other factors within the region. Complexity is also underlined as the same factor may have different effects in different regions. Education reduces prevalence in some regions but may also increase it by stimulating economic development and a wealthier society. In Iringa, high levels of education, contributing to wealth, exacerbates the prevalence. In Lindi, education reduces the presence of HIV.

These complex relationships and the tendency for factors to influence prevalence both positively and negatively obscure correlations. As a result, statistical tests give minimal information on the relationship between HIV prevalence and the factors controlling it (see Appendix III), highlighting the complex relationships.

Figure 13: Graph to show Spearman's Rank Correlation Coefficient



(Emma Harrington, 2014)

The minimal correlation showed by the correlation coefficients highlight the complexity of the issue as factors influence in both directions, cancelling out any correlation. Results of statistical analysis suggest that education is a controlling

factor in more regions. However, from analysing the data using maps it is clear that where wealth is present, education is not significant in reducing HIV prevalence.

Although the data source is a reliable survey, it is likely that there are flaws in the data. Inaccuracies will be present in such a large scale project as only a small sample can be taken in comparison to the entire population size. Moreover, it is likely that the prevalence values are inaccurate as such stigma around HIV in Tanzania may cause a reluctance to disclose HIV status.

The relationship between education and wealth is important in terms of understanding the epidemic. It doesn't seem that disparity in wealth within a region drives the epidemic. In fact the most equal areas tend to be those least impacted by HIV (see appendix II). Education is undoubtedly effective to some extent but, where wealth is present, lifestyle differs so much between the rich and the poor that educational measures reduce the spread of HIV less effectively. The educational programmes provided may be more suited to the poor populations where a lack of education causes the virus to spread. However, these programmes do not target the causes of a high HIV prevalence among the wealthy- increased mobility and access to social networks and healthcare. In this way, it is not disparity in wealth within regions that drives the epidemic but is the disparity in the quality of life caused by a variation in wealth on a national scale. In a population where access to healthcare and education is not universal it is very difficult to contain the epidemic. Tanzania currently implements a universal policy to control HIV however; there is clearly large variation in the population's characteristics and in HIV prevalence. If the government were to offer a more specific, even localised policy, the overall presence of HIV could potentially be reduced.

Appendix I

Tanzania Regions 2011

Tanzania Regions 2012

In 2012, Tanzania's regions underwent a number of changes, splitting regions to create new ones. Some of the data used was based on the 2012 regions and some was based on the regions prior to the changes. For consistency, all the data has been converted to suit the earlier regions.

The previous Shinyanga region is now comprised of Shinyanga, Simyu and Geita. The values for each of these regions were averaged to form one figure for the Shinyanga region. This was also done for Rukwa (Rukwa and Katavi) and for Iringa (Njombe and Iringa).

Zanzibar has not been included in the data.

Appendix II

Region	Total HIV positive (%)
Dodoma	2.9
Arusha	3.2
Kilimanjaro	3.8
Tanga	2.4
Morogoro	3.8
Pwani	5.9
Dar es Salaam	6.9
Lindi	2.9
Mtwara	4.1
Ruvuma	7.0
Iringa	12.0
Mbeya	9.0
Singida	3.3
Tabora	5.1
Rukwa	6.1
Kigoma	3.4
Shinyanga	5.2
Kagera	4.8
Mwanza	4.2
Mara	4.5
Manyara	1.5

Table to show Total Proportion HIV Positive by region

Table to show Total Proportion with Primary Education by region

Region	Total Primary Education (%)
Dodoma	74.7
Arusha	65.1
Kilimanjaro	82.2
Tanga	66.8
Morogoro	74.7
Pwani	72.6
Dar es Salaam	83.4
Lindi	72.5
Mtwara	75.9
Ruvuma	78.5
Iringa	81.1
Mbeya	68.2
Singida	68.8
Tabora	61.6
Rukwa	56.2
Kigoma	69.7
Shinyanga	58.7
Kagera	72.1
Mwanza	68.3
Mara	67.4
Manyara	71.9

Region	Population Below Poverty Line (%)
Dodoma	32
Arusha	23
Kilimanjaro	28
Tanga	27
Morogoro	26
Pwani	40
Dar es Salaam	20
Lindi	36
Mtwara	38
Ruvuma	39
Iringa	26
Mbeya	25
Singida	54
Tabora	40
Rukwa	36
Kigoma	36
Shinyanga	42
Kagera	27
Mwanza	36
Mara	53
Manyara	40

Table to show Total Proportion of Population below the Poverty Line by region [4]

Table to show the Gini Coefficient by region

Region	Gini Coefficient
Dodoma	0.53
Arusha	0.49
Kilimanjaro	0.38
Tanga	0.51
Morogoro	0.52
Pwani	0.42
Dar es Salaam	0.15
Lindi	0.50
Mtwara	0.46
Ruvuma	0.37
Iringa	0.34
Mbeya	0.44
Singida	0.60
Tabora	0.44
Rukwa	0.49
Kigoma	0.51
Shinyanga	0.44
Kagera	0.46
Mwanza	0.59
Mara	0.53
Manyara	0.54

[4]The Proportion of the Population below the Poverty Line- "The incidence of poverty is the percentage of the population whose per capita consumption
is below the poverty line, that is, the population that cannot afford the basic basket of goods (This comprises all consumption including that which is not bought on the market but produced for own consumption)." (Tanzania Poverty and Human Development Report, 2005)

Table to show the characteristics of populations by Education Level

Proportion of men with 2+ Wives

Proportion of people accepting in all 4 factors[5]

Education Level	Men with 2+ wives (%)	People Accepting in all 4 factors (%)
No education	16.9	16.9
Primary incomplete	13.2	13.2
Primary complete	10.5	10.5
Secondary+	6.3	6.3

[5] Members of the population were asked about 4 indicators and were recorded as expressing an accepting opinion or the contrary. The respondents were asked if they: are willing to take care of a family member with the AIDS virus in the home, would buy fresh vegetables from a shopkeeper with the virus, think a female teacher who is HIV positive but not sick should continue teaching, would not keep it secret if a family member was infecting with HIV.

Appendix III

Spearman's Rank Test

Table to show Spearman's Rank Test of Percentage of Population with Population below Poverty Line vs. HIV Prevalence

Region	Data 1	Data 2	Rank 1	Rank 2	d	d^2
Manyara	1.5	40	1.0	17.0	16	256
Tanga	2.4	27	2.0	6.5	4.5	20.25
Dodoma	2.9	32	3.5	9.0	5.5	30.25
Lindi	2.9	36	3.5	11.5	8	64
Arusha	3.2	23	5.0	2.0	-3	9
Singida	3.3	54	6.0	21.0	15	225
Kigoma	3.4	36	7.0	11.5	4.5	20.25
Kilimanjaro	3.8	26	8.5	8.0	-0.5	0.25
Morogoro	3.8	28	8.5	4.5	-4	16
Mtwara	4.1	38	10.0	14.0	4	16
Mwanza	4.2	36	11.0	11.5	0.5	0.25
Mara	4.5	53	12.0	20.0	8	64
Kagera	4.8	27	13.0	6.5	-6.5	42.25
Tabora	5.1	40	14.0	17.0	3	9
Shinyanga	5.2	42	15.0	19.0	4	16
Pwani	5.9	40	16.0	17.0	1	1
Rukwa	6.1	36	17.0	11.5	-5.5	30.25
Dar es Salaam	6.9	20	18.0	1.0	-17	289
Ruvuma	7.0	39	19.0	15.0	-4	16
Mbeya	9.0	25	20.0	3.0	-17	289
Iringa	12.0	26	21.0	4.5	-16.5	272.25
					TOTAL	1686
					P	-0.09

Spearman's Rank Test

Table to show Spearman's Rank Test of Percentage of Population with Primary Education (minus extremes) vs. HIV Prevalence

(Excludes Dar es Salaam, Iringa, Arusha, Ruvuma)

Region	Data 1	Data 2	Rank 1	Rank 2	d	d^2
Manyara	1.5	71.9	1.0	10.0	9.0	81.0
Tanga	2.4	66.8	2.0	4.0	2.0	4.0
Lindi	2.9	72.5	3.5	12.0	8.5	72.3
Dodoma	2.9	74.7	3.5	15.0	11.5	132.3
Singida	3.3	68.8	5.0	8.0	3.0	9.0
Kigoma	3.4	69.7	6.0	9.0	3.0	9.0
Morogoro	3.8	74.7	7.5	14.0	6.5	42.3
Kilimanjaro	3.8	82.2	7.5	18.0	10.5	110.3
Mtwara	4.1	75.9	9.0	16.0	7.0	49.0
Mwanza	4.2	68.3	10.0	7.0	-3.0	9.0
Mara	4.5	67.4	11.0	5.0	-6.0	36.0
Kagera	4.8	72.1	12.0	11.0	-1.0	1.0
Tabora	5.1	61.6	13.0	3.0	-10.0	100.0
Shinyanga	5.2	58.7	14.0	2.0	-12.0	144.0
Pwani	5.9	72.6	15.0	13.0	-2.0	4.0
Rukwa	6.1	56.2	16.0	1.0	-15.0	225.0
Ruvuma	7.0	78.5	17.0	17.0	0.0	0.0
Mbeya	9.0	68.2	18.0	6.0	-12.0	144.0
					TOTAL	1172.0
					P	-0.21

Spearman's Rank Test

Table to show Spearman's Rank Test of Percentage of Population with Primary Education vs. HIV Prevalence

Region	Data 1	Data 2	Rank 1	Rank 2	d	d^2
Manyara	1.5	71.9	1.0	11.0	10.0	100.0
Tanga	2.4	66.8	2.0	5.0	3.0	9.0
Dodoma	2.9	74.7	3.5	13.0	9.5	90.3
Lindi	2.9	72.5	3.5	16.0	12.5	156.3
Arusha	3.2	65.1	5.0	4.0	-1.0	1.0
Singida	3.3	68.8	6.0	9.0	3.0	9.0
Kigoma	3.4	69.7	7.0	10.0	3.0	9.0
Kilimanjaro	3.8	82.2	8.5	15.0	6.5	42.3
Morogoro	3.8	74.7	8.5	20.0	11.5	132.3
Mtwara	4.1	75.9	10.0	17.0	7.0	49.0
Mwanza	4.2	68.3	11.0	8.0	-3.0	9.0
Mara	4.5	67.4	12.0	6.0	-6.0	36.0
Kagera	4.8	72.1	13.0	12.0	-1.0	1.0
Tabora	5.1	61.6	14.0	3.0	-11.0	121.0
Shinyanga	5.2	58.7	15.0	2.0	-13.0	169.0
Pwani	5.9	72.6	16.0	14.0	-2.0	4.0
Rukwa	6.1	56.2	17.0	1.0	-16.0	256.0
Dar es Salaam	6.9	83.4	18.0	21.0	3.0	9.0
Ruvuma	7.0	78.5	19.0	18.0	-1.0	1.0
Mbeya	9.0	68.2	20.0	7.0	-13.0	169.0
Iringa	12.0	81.1	21.0	19.0	-2.0	4.0
					TOTAL	1377.0
					P	0.106

Bibliography

AIDS (06.06.2012) *Global Statistics* [Online] Available: http://www.aids.gov/hiv-aids-basics/hiv-aids-101/global-statistics/ [Accessed: 12th June 2014]

AIDS (19.12.2013) *Stages of HIV Infection* [Online] Available: http://www.aids.gov/hiv-aids-basics/just-diagnosed-with-hiv-aids/hiv-in-your-body/stages-of-hiv [Accessed: 19th July 2014]

AIDS info (17.12.2013) *Online Database* [Online] Available: http://www.aidsinfoonline.org/devinfo/libraries/aspx/Home.aspx [Accessed 12th June 2014]

All Africa (23.09.2012) *Tanzania: Mining at 4.6 Percent of GDP* [Online] Available: http://allafrica.com/stories/201209250228.html [Accessed: 24th July 2014]

AVERT (2012) *HIV & AIDS in Tanzania* [Online] Available: http://www.avert.org/hiv-aids-tanzania.htm [Accessed: 12th June 2014]

CIA (2012) *The World FactBook* [Online] Available: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2155rank.html [Accessed: 12th June 2014]

GREENWOOD D. et al (eds.) (2007) *Medical Microbiology.* 17th Edition. Elsevier Limited.

http://caitfish88.wordpress.com/2012/03/02/hiv-rate-of-lesotho-is-24-the-third-highest-in-the-world/ [Accessed: 18th August 2014]

Mining Weekly (2005) *Bulyanhulu Mine, Tanzania* [Online] Available: http://www.miningweekly.com/article/bulyanhulu-mine-2005-10-07 [Accessed: 23rd September 2014]

PRB (2012) *World Population Data Sheet 2012* [Online] Available: http://www.prb.org/Publications/Datasheets/2012/world-population-data-sheet/data-sheet.aspx [Accessed: 12th June 2014]

SHELTON, J. et al (2005) Is Poverty of Wealth at the root of HIV? *The Lancet* 366 (September) p. 1057-1058

TAN, J. (2001) *Expert Guide to Infectious Diseases*. Expert Guide Series

Tanzania Commission for HIV/AIDS (2003) *National Multi-Sectoral Strategic Framework on HIV/AIDS 2003 - 2007.* [Online] Ifakara Health Institute. Available: http://ihi.eprints.org/640/ [Accessed: 12th June 2014]

Tanzania Ports Authority (2012) *Dar Es Salaam Port* [Online] Available: http://www.tanzaniaports.com/index.php?option=com_content&view=article&id=100&Itemid=270 [Accessed: 19th August 2014]

Tanzania Poverty and Human Development Report (2005) *Poverty Line* [Online] Available: http://www.repoa.or.tz/documents/PHDR_2005_Prelim.pdf [Accessed: March 2014]

The World Bank (2012) *Data* [Online] Available: http://data.worldbank.org/region/sub-saharan-africa [Accessed: 12th June 2014]

Welch, J. (2014) *HIV Clinic* Interviewed by: Harrington, E. Kings College Hospital, London

WHITESIDE, A. (2008) *HIV/AIDS A Very Short Introduction.* Oxford: Oxford University Press

Wikipedia (23.07.2011) *Regions of Tanzania* [Online] Available: http://en.wikipedia.org/wiki/File:Regions_of_Tanzania_sw.svg [Accessed: 20th August 2014]

How is the war of 1812 between the United States and Britain represented in art and literature by each side during the conflict?

Claudia Hockey - History

The Extended Essay is a great opportunity for any historian to research and produce a unique piece of work, irrespective of topics covered on either of the HL History routes. Although my course centres on aspects of Medieval Europe and the Islamic World, this essay gave me the chance to explore freely my interest in 18th and 19th century America. The ability to sift through and select from vast quantities of increasingly digitised source material, as well as restructuring arguments to fit tight word limits are essential skills I've learnt that will hopefully stand me in good stead for university. I would advise students considering History to start thinking about topics to write about as early as possible, to leave plenty of time for research.

Supervisor: David Hall

In studying the War of 1812, Claudia was keen to avoid the almost routine focus of essays on causes and/or consequences. Her decision to focus on propaganda emerged from our discussions about possible new angles, and it drew on her interest in (and perceptive approach regarding) art and literature as well as History. In addition to the resourcefulness she showed in finding evidence, she asked very interesting questions about the possibility of artists on opposing sides actually influencing each other's work during the conflict. This led to a completely new angle in her comparative study of the propaganda, as she studied the cartoons for evidence of direct transatlantic influence.

Contents

Abstract

"How is the war of 1812 between the United States and Britain represented in Art and literature by each side during the conflict?"
The focus of this investigation is to explore, through the study of literary and visual material, the development of British and American governmental perspectives throughout the War of 1812. Within the context of the War of 1812 underlying weaknesses in Madison's leadership were revealed, immortalized in the work of cartoonists such as George Cruikshank. Examining the period between 1st June 1812 and December 1815 this essay incorporates academic research such as "The War of 1812 in the age of Napoleon" by Jeremy Black[2]. The essay is broken down into 10 sections, each outlining a piece of Art or literature and illuminating their importance in the context of both the War of 1812 and the development of Anglo-American relations as a whole. It investigates the extent to which personal motives impacted Propaganda, and how the self- representation of the powers in these pieces were indicative of an imbalance of power between the sides. By doing so it attempts to ascertain the degree to which these pieces reflect the development of American Identity in its evolution as a young republic and the role of the War of 1812 in Britain as an aspect of the larger Napoleonic Wars. Representations of the War of 1812 ultimately served to justify rather than advocate enlistment in the War, Madison's call to arms having been primarily a result of his desire to succeed in the light of successes enjoyed by Washington, Adams and Jefferson before him. The War of 1812 would set a precedent for not only Anglo-American relations, but also those between the United States and Canada, and marked the end of any chance for an independent Native American nation.

[2] Jeremy Black, The War of 1812 in the age of Napoleon (Norman, 2009)

[1]

[1] William Charles, *The Hartford Convention or Leap no leap* (Philadelphia, 1814)

Introduction

The War of 1812, despite its significance in the larger scheme of Anglo-American politics, is a relatively forgotten conflict. Part of its appeal as a historical topic is that debate stemming from its ambiguity still rages today. Viewed south of the Canadian-American border as a "nightmare from the nations childhood", northwards it's seen as a defining moment in the establishment of independence from America[3]. The sense of incoherence that surrounds the character of the War of 1812 as a conflict can be attributed to Adams' work in the 1890s, whose view that the war had begun without either side having a clear idea of its causes marked a clear contrast with those of Roosevelt and Mahan[4]. To look deeper into the ways in which the War of 1812 was represented by each side allows us to gain a more complete understanding of its significance in the context of the power struggle that had characterised Anglo-American relations up until that point, and its legacy today. The imbalance between Britain, a great imperial power, and America, a young nation with great imperial potential but in danger of disappearing as a nation state before that potential was realised, is reflected in contemporary propaganda.

The extent to which this imbalance played a role in the outcome of the war is one of the first issues I wanted to explore. However in doing so important questions are raised such as 'does the public get what it wants or want what it gets?' The motives for involvement in propaganda by both the government and the public were more complex than I had initially thought, however in this context the answer appears to lean towards the former. It's also important to consider when relevant cartoonists and their sources. A good example of the effect of external influences on propaganda can be found in the work of David Low, who in 1927 joined the Evening Standard on the understanding that there would be no editorial intervention in his work[5]. When it became clear that his political cartoons were damaging Anglo-German relations Low was forced to implant a composite dictator in place of Mussolini when producing his strip 'Hit and Muss'[6]. The question of how far the 'free' media is influenced by the government continues to hold sway today, as an individualist nation we consider the questioning of persuasive works essential. In the case of the War of 1812, there was great pressure on Madison to succeed in this conflict following the successes of Washington, Adams and Jefferson so it would have been in his interests to encourage positive depictions of the war. This essay argues that representations of the War of 1812 from both Britain and America were not a culmination of discord resulting from the War of Independence, but rather served to justify involvement in a conflict that was misunderstood by almost all who fought it.

[3] J.C.A Stagg, *The War of 1812: Conflict for a Continent (*New York, 2012), p.ix
[4] J.C.A Stagg, *The War of 1812: Conflict for a Continent (*New York, 2012), p.4
[5] Biography: David Low, *The British Cartoon Archive*
[6] Victor S. Navasky, *The Art of Controversy: Political Cartoons and Their Enduring Power* (New York, 2013), p.115

Investigation

I – James Madison Proclaims War, June 19th 1812

President Madison initially presented the War of 1812 as a means to maintain the rights achieved by the 'virtue and valour of [revolutionaries]'[7] just 36 years earlier. Madison looked to exploit a deep-set desire within the American people to uphold national honour in the face of British slights such as the Chesapeake-Leopard affair[8] under Jefferson's government:

I do therefore exhort all the good people of the United States as they love their country, as they value the precious heritage derived from the virtue and valour of their fathers, as they feel the wrongs which have forced on them the last resort of injured nations, and as they consult the best means, under the blessing of divine providence, of abridging its calamities, that they exert themselves in preserving order, in promoting concord, in maintaining the authority and the efficacy of the laws, and in supporting and invigorating all the measures which may be adopted by the constituted authorities for obtaining a speedy, a just and honourable peace'[9]

His reference to the 'wrongs' of the British is a response to their conscription of American sailors[10] and the enforcement of a blockade preventing American trade with France in response to the Napoleonic Wars. Discord stemmed largely from Britain not recognising the right of its subjects to renounce their citizenship[11]. This brought anti-British feelings to the surface among Americans and served as a launching-post off which Madison could embark on a campaign against Britain. Madison's proclamation also illustrated the importance of a War in preserving the tenuous order, concord and efficacy of laws in a newly emerging nation, to quickly achieve a 'just and honourable peace'[12]. By presenting the War of 1812 as nobly led, Madison hoped to appeal to the patriotism of the descendants of revolutionaries, his description of it as 'The last resort of injured nations' placing the US firmly amongst other nations forced into war.

This piece is particularly relevant to the question 'does man influence the media, or is the media influenced by man?' as in the years prior to this proclamation some still felt that social improvement had not been sufficient. The phraseology used by Madison e.g. "preserving order", "calamities", indicates that he recognised that war could speed up social change, and feared the young radicals and his waning control over them - the proclamation served as a subtle warning not to challenge the constitution. Moreover Madison's added reminder that 'all persons holding Offices, Civil or Military' had a duty to their country serves to directly address the sense of responsibility within his audience, words such as 'zealous' being especially effective due to their religious undertones. By drawing on themes of authority and social obligation Madison's Proclamation served to both justify and advocate the War of 1812.

---

[7] Donald R. Hickey, *The War of 1812: Writings from America's Second War of Independence* (New York, 2013), p.36
[8] Jeremy Black, *The War of 1812 in the Age of Napoleon (*Norman, 2009) p.26
[9] Donald R. Hickey, *The War of 1812: Writings from America's Second War of Independence* (New York, 2013), p.36
[10] ibid.
[11] ibid.
[12] Donald R. Hickey, *The War of 1812: Writings from America's Second War of Independence* (New York, 2013), p.36

II – James Madison War message to Congress, 1st June 1812

Despite Madison's later assertion to the American people that he looked to find a 'speedy, just and honourable peace'[13], a look back at his War Message to Congress clearly indicates that he had pressed for war. Previously, Madison had expressed scant regard for the idea that diplomacy alone could prompt the British to repeal their restrictive 'Orders in Council'.[14] Recommending that the US take "an armour and an attitude demanded by the crisis", Madison faced increasing pressure from 'War Hawks' such as Henry Clay and other Republicans of the Twelfth Congress to propose war[15]:

Against this crying enormity, which Great Britain would be so prompt to avenge if committed against herself, the United States have in vain exhausted remonstrances and expostulations, and that no proof might be wanting of their conciliatory dispositions, and no pretext left for a continuance of the practice, the British Government was formally assured of the readiness of the United States to enter into arrangements such as could not be rejected if the recovery of British subjects were the real and the sole object. The communication passed without effect.[16]

Madison's message to Congress highlighted transgressions committed by the British, who had 'wantonly spilt American blood within … [the US] territorial jurisdiction', with the intention of urging Congress into war. Words such as 'plundered', 'destructive' and 'belligerent' present Britain as predatory, an obstacle to US 'agricultural and maritime interests' and responsible for '[cutting off America] from their legitimate markets'. By portraying Britain this way Madison strengthened growing fears among Americans, Congress included, that the 'commerce of the United States [was] to be sacrificed'. In Madison's eyes the survival of the United States was entirely dependent on its success in international trade. The British blockades were purely a means to prevent American intervention with the 'monopoly which she covets for her own commerce and navigation' - Madison believed the United States was a victim of British double standards; 'which … Britain would be so prompt to avenge if committed against herself'. Madison's message also touches upon the relationship between Britain and 'the savages':

In reviewing the conduct of Great Britain toward the United States our attention is necessarily drawn to the warfare just renewed by the savages on one of our extensive frontiers—a warfare which is known to spare neither age nor sex and to be distinguished by features peculiarly shocking to humanity.[17]

American-Indian conflict, sporadic since colonists first began settling in Northern America, was exploited by the British in the War of 1812, who saw the tribes as a valuable buffer between their Canadian settlements and American forces. The British also felt a sense of obligation towards Native Americans following their abandonment in the Anglo-American treaties of 1783 and 1794, and increasing popularity had arisen for the idea of the 'Noble Savage', spurred by works such as *'Hermsprong: or Man as He is Not'* by Robert Bage.[18] It was no great secret that America blamed the British for increased native resistance led by Tecumseh and Tenskwatawa, "the Prophet", and American suspicions over British-Indian relations had led to a desire to seize Canada, an area seen as the primary link between the two parties.[19] By emphasising British involvement in the American struggle with 'savages' Madison successfully drew together pre-existent fears regarding Native American brutality and lingering distaste towards British forces following the War of Independence.

---

[13] Donald R. Hickey, *The War of 1812: Writings from America's Second War of Independence* (New York, 2013), p.36
[14] Donald R. Hickey, *The War of 1812: Writings from America's Second War of Independence* (New York, 2013), p.1
[15] ibid.
[16] Donald R. Hickey, *The War of 1812: Writings from America's Second War of Independence* (New York, 2013), p.2
[17] Donald R. Hickey, *The War of 1812: Writings from America's Second War of Independence* (New York, 2013), p.7
[18] Jeremy Black, *The War of 1812 in the Age of Napoleon (*Norman, 2009) p.40
[19] Jeremy Black, *The War of 1812 in the Age of Napoleon (*Norman, 2009) p.41

As in his address to the people 18 days later Madison was careful in this message to strike a balance between the presentation of the United States as ‘an injured nation’ attempting to avenge itself against a ‘crying enormity’, and a ‘free and powerful’ country, driven by patriotic fervour. However despite the implications of Madison’s message he makes no specific declaration of war, instead leaving the question open to the ‘legislative department of the Government’, a humble stance markedly different to his later approach.

III – The 'Humane' British and their worthy allies

20

The above 1812 cartoon by William Charles encapsulates the moral attitude of the American people upon the war's outbreak. The cartoon may have been a response to the Chicago massacre in August 1812, where Native Americans scalped American officers and ate their hearts.[21] The linked hands of both British soldier and Indian over an American scalp emphasizes that their union is forged in violence and sealed with American blood. Rather than allying with another Christian country, the British have made a deal with the devil's instruments. Charles ignores the fact that Americans had also allied with Indians[22], instead using forest imagery to highlight the fear of the unknown that accompanied American contemporary opinion of 'savages'. Later, Robert Frost's 'Stopping by woods on a snowy evening'[23] highlights a similar idea, that a forest represents an area exempt from societal values.

The metaphorical significance of a forest infested with Indians, primitive and beyond the reach of societal influence, is exploited here to criticize intervention by the 'Humane British' as being the lowest form of moral conduct. Charles' use of colour, green for the Indians as archaic beings at one with nature, blood red for the British as those responsible for the war and depravity, serves to highlight the status of American soldiers as victims – an idea previously encouraged by Madison in his justification for war[24]. Two American soldiers dressed in blue represent purity in the face of this corruption and

---

[20] William Charles, *A scene on the frontiers as practiced by the "humane" British and their "worthy" allies* (Philadelphia, 1812)
[21] Milo Milton Quaife, *Chicago and the Old Northwest, 1673-1835; a study of the evolution of the northwestern frontier, together with a history of Fort Dearborn* (Chicago, 1913) p.411
[22] Thomas Jefferson, *To the Brothers of the Choctaw Nation,* Yale Law School (2008)
[23] Robert Frost, *Stopping by Woods on a Snowy Evening*, The Poetry Foundation, first published New Hampshire Volume (1923)
[24] Donald R. Hickey, *The War of 1812: Writings from America's Second War of Independence* (New York, 2013), p.36

sin, the Indian soldiers scalping them carrying British guns labeled 'reward for sixteen scalps'. The knives at their waists are festooned with the symbol 'GR', an abbreviation of 'King George' and thus an explicit reference to Royal involvement. Charles is drawing on the popular idea that Britain was a nation founded in degeneracy and greed. This is further underlined in the red-face of the British soldier, he is portrayed as a drunkard fond of over-eating at the expense of lesser nations. The British are represented as morally and physically repulsive, the verse at the bottom of the cartoon giving the impression that these scalpings are happening on a large scale:

Arise Columbia's Sons and forward press, / Your Country's wrongs call loudly for redress; / The Savage Indian with his Scalping knife, / Or Tomahawk may seek to take your life; / By bravery aw'd they'll in a dreadful Fright, / Shrink back for Refuge to the Woods in Flight; / Their British leaders then will quickly shake, / And for those wrongs shall restitution make.[25]

The Irony within the poem's title and the wholly negative depiction of British forces may be a reflection not only of adverse public opinion, the cartoon being released to an American audience, but also the experiences of the cartoonist. Scottish-born, little is known about his early life but he may have immigrated to the United States following backlash from his satirical cartoons, having worked in London and Edinburgh previously[26].

[25] William Charles, *A scene on the frontiers as practiced by the "humane" British and their "worthy" allies* (Philadelphia, 1812)
[26] Lorraine Lanmon, American Caricature in the English Tradition: The Personal and Political Satires of William Charles, *The University of Chicago Press,* 11, (1976) p.1

IV – A sketch for the Regent's speech on 'Mad-ass-son's' insanity

The British cartoonist George Cruikshank's 1812 work, 'A sketch for the regents speech on Mad-ass-son's insanity' strikes an entirely different tone in its depiction of early Anglo-American conduct.

[27]

Rather than presenting Britain as a victim, the cartoon emphasises the foolhardiness of Madison in challenging the British Empire. Napoleon's presence indicates that they are not alone in this opinion: other great nations also look down on the United States. An angelic figure adds to the ridicule, chorusing 'A bad news for you' at Madison, who stands between Napoleon and the Devil, as two women symbolising Britain and America, along with soldiers, look on. The President, a laughing stock even in name, 'Mad-ass-son', has been firmly separated from his nation; the blame here is not being placed on Americans as a body but on Madison as their leader. The 'bad news' in question refers to the defeat of General Hull at the siege of Detroit,[28] who was intimidated by the British into surrendering the fort and town following reports that there were far larger numbers of Native Americans than were actually present.[29] Bold colours combined with the presence of British troops reinforce the idea that Britain was well supported compared to the solitary 'Mad-ass-son'. The lush British grass, in comparison to the sparse, desert like conditions beneath Madison, the devil and Napoleon marks a clear distinction between good and bad parties. The Lion, reiterating British pride for its heritage, is worthy of note given Cruikshank's later notoriety following his acceptance of a royal bribe of £100 for a pledge 'not to caricature his majesty...in any immoral situation'.[30]

[27] George Cruikshank, *A sketch for the regents speech on Mad-ass-son's insanity* (London, 1812)
[28] Andrew Lambert, *The Challenge: America, Britain and the War of 1812* (London, 2012), p.133
[29] Jeremy Black, *The War of 1812 in the Age of Napoleon (*Norman, 2009) p.64
[30] George Cruikshank: A Successor In Satire, *Newcastle University Library* (2014)

V – 'The Yankey Torpedo'

[31]

The work of British-born Thomas Tegg a year later demonstrates developing attitudes among the British towards American proficiency in war. Rather than ridiculing Madison's stance, the piece has a resentful atmosphere at American usage of the torpedo similar to that of Charles towards the British in Illustration III[32]. However whilst in this piece John Bull, a personification of the British people, initially resembles a victim of enemy tactics as American soldiers were in Charles' work, his expression and upright stance indicate righteous indignation rather than a desperate appeal to arms. The demonic imagery used, emerging from the depths of hell to pollute British air, reflects British distaste towards Americans as in Cruikshank's 1812 cartoon. The positioning of John Bull, away from the troubles posed by North America, is suggestive of British preoccupation with the larger Napoleonic Wars. Although the use of torpedoes by Americans following the passing of the 'Torpedo Act', which stated that anyone who sank a British ship using one was eligible to half its value[33], did not lead to naval success in the War of 1812, it contributed to the establishment of the United States as a global naval power - the navy becoming preferable to the army among many American citizens.[34]

[31] Thomas Tegg, *The Yankey Torpedo* (London, 1813)
[32] William Charles, *A scene on the frontiers as practiced by the "humane" British and their "worthy" allies* (Philadelphia, 1812)
[33] Jeremy Black, *The War of 1812 in the Age of Napoleon (*Norman, 2009) p.145
[34] ibid.

VI – Shannon versus Chesapeake

35

A sense of superiority among the British continued through into September of 1813, Cruikshank's work portraying Americans as a subject of ridicule much like his previous portrayal of 'Mad-ass-son'. British troops aboard the Chesapeake call out for American soldiers to return to their dinner, an allusion to the United States being in a state of infancy. This contrasts with the bravado of British troops, whose well-turned out uniforms are juxtaposed with the assortment of mismatched clothing worn by Americans. This British sense of unity is also highlighted in their positioning; with grand postures they are depicted as moving as one unit, as opposed to Cruikshank's submissive and disorganized American troops. Representations of the opposing sides at this time vary in that American propaganda never depicts the British as lying on the floor, they remain in fighting positions even when failing, whilst soldiers from the United States are consistently illustrated in vulnerable positions befitting their attitude as victims of British slights.

35 George Cruikshank, *British valour and Yankee boasting or, Shannon versus Chesapeake* (unknown publisher, 1813)

VII – Boxing match, or another bloody nose for John Bull?

Published in the same year as 'Shannon versus Chesapeake' Charles' work instead highlights American valor in another naval victory, the defeat of the British 'Boxer' by the warship 'Enterprise' in September 1813[36].

37

Although the sea in the background of this piece appears empty, in reality the British Army boasted 10,000 disciplined regulars, compared to around 7,000 Americans[38]. They also had an advantage in their alliance with Indians, American militants being forced to wage defensive warfare on a frontier with crude roads, unreliable waterways[39] and little food. These difficulties are overlooked by Charles, as is the involvement of Canadian and Indian militants, reflecting the United States' opinion that the War of 1812 was very much Anglo-American in character. The opulent clothing of George III emphasizes again the poetic justice of a nation that's 'underdressed' in comparison coming out on top, brash red has been surpassed by a more subdued black. His broken and crooked crown caters to anti-monarchist feelings among Americans, the wounded pride of a British monarch appealing to those hoping to preserve their Nation's democratic integrity. The title of this piece carries the implication that the conflict can hardly qualify as a 'match', the imbalance between the two being addressed such that America appears to have the upper hand when in reality Charles was hyperbolizing minor victories.

---

36 Barry J. Lohnes, British Naval Problems at Halifax During the War of 1812, *Mariner's Mirror,* 59 (1973) p.326

37 William Charles, A boxing match, or another bloody nose for John Bull? (unknown publisher, 1813)

38 Stuart Murray, *Atlas of American Military History,* (New York, 2005) p.53

39 Donald R. Hickey, *The War of 1812: A Short History,* (Champaign, 1995) p.23

VIII – The fall of Washington

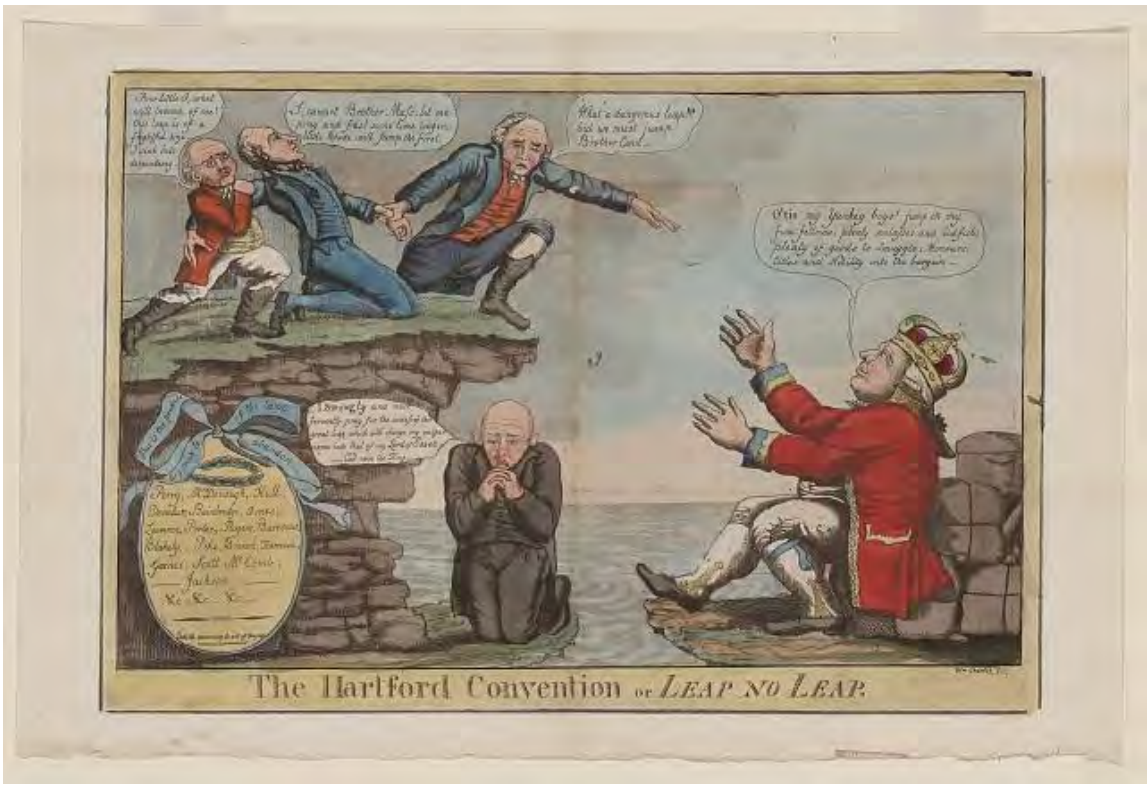
This piece carries a similar gloating tone to Charles' work, perhaps as a response to American enthusiasm the prior year. Although the artist is unknown, judging by the style we can safely assume that it is Cruikshank.

40

The 1814 cartoon depicts Madison and perhaps John Armstrong, his Secretary of War, fleeing from Washington with papers tucked under their arms. This picture of disarray reflects the poor preparation of Madison and his party for an invasion by the British in retaliation for the destruction at York in 1813. This was in part due to the unfavourable odds facing British troops, the approach to shore and disembarkation were slow and complex, and at Bladensburg on 24th August the British found themselves without cavalry and heavily outgunned, up against sixty-five hundred Americans. The account of American citizen John Threlkeld scathingly notes the inadequacy of United States militia, "Perhaps none ever saw a gun pointed at a man to kill in their lives". Perhaps the most crucial factor in the fall of Washington lies in the attitude of the local populace, illustrated here. In the defeat, as at Boston, New York and Philadelphia in the late 1700s, resistance to British forces remained exclusively beyond city limits, with little insurrection inside. American leadership in the aftermath of the disaster was divided, Armstrong becoming an opportune scapegoat following complaints among militia regarding his inaction.[41]

[40] Unknown, *The fall of Washington, or Maddy in full flight,* (London, 1814)
[41] Jeremy Black, *The War of 1812 in the Age of Napoleon (*Norman, 2009) pp. 170-175

IX – The Hartford Convention



42

Animosity towards American leadership culminated in the Hartford Convention. Illustrated here by Charles, whose previous works were firmly supportive of US high powers, the convention was a series of secret meetings set up by Federalists who strongly opposed the conflict. Whilst outwardly the convention would discuss local defense and taxation, it also posed an opportunity for Federalists to air their grievances with Republican leadership[43]. Initially Madison ignored dissent on the home front, however Charles points out the dire consequences of radical secessionist action by depicting King George as opportunistic in the face of American discord. Charles once again emphasizes a basic fear of the unknown, (as in Illustration III), in his portrayal of convention members, who, surrounded by seawater, jump straight into the hands of a predator. Barren waters behind them highlight the isolation of the United States, the boxes beneath George III representing goods denied to them due to trade blockades. The sharp contrast between the submissive positioning of Timothy Pickering on the underside of a cliff, 'I.. pray for the success of this great leap which will change my vulgar name', and King George, 'jump in my fine fellows…Honours, titles and nobility', whose arms are outstretched lends further emphasis to the idea that by doing this convention members would be playing right into the hands of King George. Furthermore a medallion on the left side of the cartoon is inscribed with the names of Perry, McDonough, Hull and other heroes of the War of 1812 and is decorated with a ribbon reading, 'This is the produce of the land they wish to abandon'. Whilst the Americans have once again been depicted as victims, much the same as in 'Shannon versus Chesapeake' and other works, the image has progressed to suggest that they are willful in this – they are jumping into the mouth of the monster below, rather than turning away from it as in 'the Yankey torpedo' cartoon.

---

42 William Charles, *The Hartford Convention or Leap no leap* (Philadelphia, 1814)
43 J.C.A Stagg, *The War of 1812: Conflict for a Continent (*New York, 2012), p.147

X – James Madison special message to Congress, February 1815

The end of the War of 1812 was not due to internal discord, nor to a resounding victory from either party, but rather a mutual understanding of the negative effects a conflict this expensive was having on both sides. The first Paris treaty marking the end of the Napoleonic Wars on May 30th 1814 left conflict-weary British citizens seeking a means to end the Anglo-American war that continued to drain their resources. The United States, heavily in debt and faced with complaints from militia who had lost sight of why they were fighting, were equally desperate to find a quick resolution. This led to the signing of the Treaty of Ghent on Christmas Eve 1814.[44] The treaty, whilst this time offering some protection to 'the Tribes or Nations of Indians', mentions nothing of the key US motivations for going to war - there is no referral to conscription of soldiers, trade embargoes nor American rights and independence. This contrasts sharply with the impression Madison gives in his message to Congress the following February, which reads as an attempt to justify his decision to go to war in the first place, given that many issues such as the British Orders in Council were repealed before the conflict even began:[45]

I congratulate you and our constituents upon an event which is highly honourable to the nation, and terminates with peculiar felicity a campaign signalized by the most brilliant successes. The late war, although reluctantly declared by Congress, had become a necessary resort to assert the rights and independence of the nation. It has been waged with a success which is the natural result of the wisdom of the legislative councils, of the patriotism of the people, of the public spirit of the militia, and of the valour of the military and naval forces of the country Peace, at all times a blessing, is peculiarly welcome, therefore, at a period when the causes for the war have ceased to operate, when the Government has demonstrated the efficiency of its powers of defence, and when the nation can review its conduct without regret and without reproach.[46]

Although the 'just and honourable peace' Madison spoke of in his 1812 Proclamation had been reached, it was by no means 'speedy' or without complications. The cartoon 'Maddy in full flight' serves as a stark contradiction to Madison's claims that 'Government [had] demonstrated the efficiency of its powers of defence', and the successes he emphasises in this piece were in reality sporadic, the 'rights and independence of the nation' – from an external viewpoint at least - largely unchanged. However the 'public spirit' that developed as a result of this conflict would remain an integral part of the American national character for years to come. This is a theme addressed by Napoleon who stated, "the Americans…wanted a ten years war to make them a nation." His comments on the resolve of the United States to place itself firmly into the context of the great European Naval powers would foreshadow future events, however the War of 1812 marked the last time Anglo-American rivalry would result in conflict – despite the United States' later surpassing the British Empire as a global power.[47] The conflict, whilst preventable in many respects, was successful in creating long-term Anglo-American concord.

[44] *Treaty of Peace and Amity between His Britannic Majesty and the United States of America*, Yale Law School (2008)
[45] Carl Benn, *Essential Histories: The War of 1812* (Oxford 2002) p.81
[46] Donald R. Hickey, *The War of 1812: Writings from America's Second War of Independence* (New York, 2013), p.689
[47] Jeremy Black, *The War of 1812 in the Age of Napoleon (*Norman, 2009) pp.239-40

Conclusion

A closer look at the propaganda produced by both Britain and America during the War of 1812 reflects the sense of ambiguity surrounding the conflict. Trivial insults are punctuated with incredulity from both sides at decisions taken by Madison, as well as the naval dimension of the conflict able to be twisted in different ways – was American involvement in this field a clear demonstration of their strength or an indicator of the need for improvement? Events during the War of 1812 would result in a measured peace between Britain and the United States that remains to this day. This can be attributed partly to the realisation by the US that Canada could not be conquered, and Britain's preoccupation with the Napoleonic wars preventing a full-scale conflict[48]. Britain's abandonment of Native Americans once their goals had been achieved proved fatal for the latter party, but was accommodating to US plans regarding westward expansion.[49] Whilst at the beginning of the conflict Madison and Cartoonists such as Charles remained adamant that British transgressions had forced their hand in declaring war, internal discord soon saw the blame being shifted to figures such as John Armstrong[50] following military incursion. Developments in American propaganda saw the portrayal of the British as autocratic and corrupt become a springboard for the emergence of a United States image equal to the British Empire in national pride if not in imperial strength. This investigation was limited by the word count restricting the range of sources used to government representation rather than the views of soldiers or civilians. Although the essay touches on later depictions of the war by figures such as Mahan no source material is analysed from the years prior to or following the conflict, preventing a deeper appraisal of its representation throughout history. To analyse British and American sources from both before and after the war would further highlight why American martial disasters in 1812 evoked in its citizens the need for military strength, namely the development of a blue-water navy and an augmented standing army in light of the United States' geographical isolation. To gain a better idea of the effect of the War of 1812 globally the viewpoint of France, responsible in part for both provoking and ending the war, would be worth looking into. From this essay it is possible to conclude that hyperbolised depictions of national strength evident in visual and literary material were just a means to conceal insecurities regarding the conflict, recurring imagery such as a fear of the unknown and Satan highlighting the indefinite nature of the war in both its causes and aims.

[48] Jeremy Black, *The War of 1812 in the Age of Napoleon (*Norman, 2009) p.240
[49] ibid.
[50] Jeremy Black, *The War of 1812 in the Age of Napoleon (*Norman, 2009) p.175

Bibliography

Primary Sources

Printed Sources:

Hickey, D.R. *The War of 1812: Writings from America's Second War of Independence.* (New York, 2013)
: James Madison: Proclamation of War, June 1812
: James Madison: War Message to Congress, June 1812
: James Madison: Special Message to Congress, February 1815

Websites:

Charles, W. *A boxing match, or another bloody nose for John Bull?* (unknown publisher, 1813) <http://www.loc.gov/pictures/item/2002708982/> [accessed at 18.40 11th August 2014]

Charles, W. *A scene on the frontiers as practiced by the "humane" British and their "worthy" allies* (Philadelphia, 1812) <http://www.loc.gov/pictures/item/2002708987/> [accessed at 17.36 2nd August 2014]

Charles, W. *The Hartford Convention or Leap no leap* (Philadelphia, 1814) <http://www.loc.gov/pictures/item/2002708988/> [accessed at 18.45 4th August 2014]

Cruikshank, G. *A sketch for the regents speech on Mad-ass-son's insanity* (London, 1812) http://www.loc.gov/pictures/item/98514002/ [accessed at 19.42 11th August 2014]

Cruikshank, G. *British valour and Yankee boasting or, Shannon versus Chesapeake* (unknown publisher, 1813) http://www.loc.gov/pictures/item/99471628/ [accessed at 19.56 9th August 2014]

Frost, R. *Stopping by Woods on a Snowy Evening*, The Poetry Foundation <http://www.poetryfoundation.org/poem/171621> [accessed at 19.21 3rd August 2014]

Jefferson, T. *To the Brothers of the Choctaw Nation,* Yale Law School (2008) <http://avalon.law.yale.edu/19th_century/jeffind3.asp> [accessed at 14.56 10th August 2014]

Tegg, T. *The Yankey Torpedo* (London, 1813) <http://www.loc.gov/pictures/item/2002720231/> [accessed at 20.03 8th August 2014]

Treaty of Peace and Amity between His Britannic Majesty and the United States of America, Yale Law School (2008) <http://avalon.law.yale.edu/19th_century/ghent.asp> [accessed at 16.21 6th August 2014]

Unknown, *The fall of Washington, or Maddy in full flight,* (London, 1814) <http://www.loc.gov/pictures/item/94507312/> [accessed at 15.44 3rd August 2014]

Secondary Sources

Printed Sources:

Benn, C. *Essential Histories: The War of 1812* (Oxford 2002)

Black, J. *The War of 1812 in the age of Napoleon* (Norman, 2009)

Lambert, A. *The Challenge: America, Britain and the War of 1812* (London, 2012)

Lohnes, B.J. British Naval Problems at Halifax During the War of 1812, *Mariner's Mirror,* 59 (1973)

Stagg, J.C.A. *The War of 1812: Conflict for a Continent (*New York, 2012)

Websites:

George Cruikshank: A Successor In Satire , *Newcastle University Library* (2014) <http://www.ncl.ac.uk/library/special-collections/exhibitions/current-and-past-exhibitions/gillray/science/cruikshank> [accessed at 19.02 7th August 2014]

Hickey, D.R. *The War of 1812: A Short History,* (Champaign, 1995) p.23 <http://books.google.co.uk/books?id=wU8LNjjfLmQC&pg=PA23&lpg=PA23&dq=war+of+1812+crude+roads+waterways&source=bl&ots=xzWLS6IrQG&sig=SME902s3A3nSvfBLKVcvA-3U1V8&hl=en&sa=X&ei=BhTyU-XBHejm7AaHiIHQBA&ved=0CCAQ6AEwAA#v=onepage&q=war%20of%201812%20crude%20roads%20waterways&f=false> [accessed at 10.31 14th August 2014]

Lanmon, L. American Caricature in the English Tradition: The Personal and Political Satires of William Charles, *The University of Chicago Press,* 11 (1976) p.1 <http://www.jstor.org/discover/10.2307/1180589?uid=3738032&uid=2&uid=4&sid=21104540236863> [accessed at 17.32 10th August 2014]

Murray, S. *Atlas of American Military History,* (New York, 2005) p.53 <http://books.google.co.uk/books?id=bJ_sy7mmmxQC&pg=PA53&lpg=PA53&dq=10,000+disciplined+regulars+in+Canada&source=bl&ots=Qh_p1bbT8K&sig=ZjmoyGwLr-fKRoIwO1KL3XoJoKg&hl=en&sa=X&ei=dhLyU8CnFqWR7Aa1g4C4Cw&ved=0CCAQ6AEwAA#v=onepage&q=10%2C000%20disciplined%20regulars%20in%20Canada&f=false> [accessed at 13.32 10th August 2014]

Navasky, V.S. *The Art of Controversy: Political Cartoons and Their Enduring Power* (New York, 2013) <http://books.google.co.uk/books?id=NBYq8e06qEcC&pg=PA113&source=gbs_toc_r&cad=3#v=onepage&q&f=false> [accessed at 13.32 7th August 2014]

Quaife, M. *Chicago and the Old Northwest, 1673-1835; a study of the evolution of the northwestern frontier, together with a history of Fort Dearborn* (Chicago, 1913) <https://archive.org/details/chicagooldnorthw00quai> [accessed at 15.32 17th August 2014]

Biography: David Low, *The British Cartoon Archive* <http://www.cartoons.ac.uk/artists/davidlow/biography> [accessed at 14.26 18th August 2014]

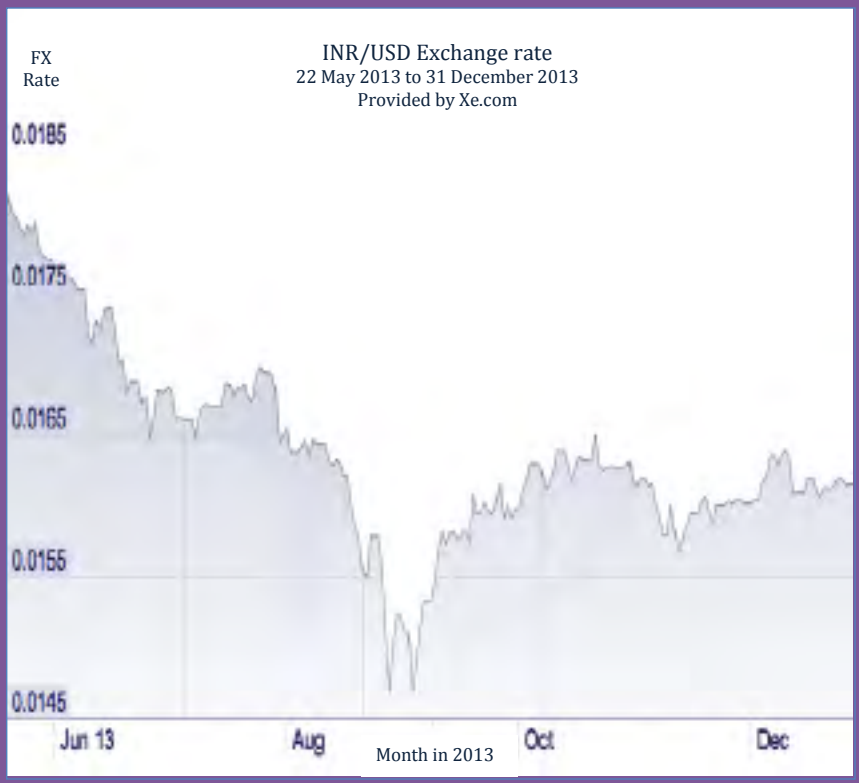
To what extent can the depreciation of the rupee against the U.S. dollar, from 22nd May 2013 to the end of 2013, be attributed to Federal Reserve policy?

Max Kitson - Economics

Approaching the task of writing a 4,000 words Extended Essay in Economics seemed daunting at first. However, having found a suitable topic, I was soon worried that 4,000 words would not be enough! For me personally, the Extended Essay served as an opportunity to explore a subject which interested me in a level of detail beyond that offered by the IB syllabus. While this was inherently enjoyable, I found that it also enhanced my university application by deepening my understanding of Economics, and by showcasing my ability to successfully produce an in-depth research project in the subject. Overall, I found the process hugely rewarding.

Supervisor: Paul Eversfield

Max was the definition of independence for his EE in that I barely saw or spoke to him: he was so engaged in the process and enjoying the research and writing that he forgot to check in half the time. What he produced though was a first rate economic analysis – a detailed application of some high level mathematical modelling to economic finance issues relating to the pricing of bonds. Clearly it is something that appeals to him and given the level of skill and flair demonstrated here, I fully expect to see him as one of the future "masters of the universe"!



Contents

Abstract

This essay considers the causes of a 10.91% depreciation of the INR/USD exchange rate from 22nd May 2013 to the end of 2013. Three main factors will be considered: the impact of the Quantitative Easing (QE) tapering announcement made on 22nd May by Chairman of the Federal Reserve Ben Bernanke, the impact of India's macroeconomic problems (specifically slowing economic growth, a large government budget deficit and high inflation) on capital flows and the impact of India's worsening trade position with the U.S.

My research question is "To what extent can the depreciation of the rupee against the U.S. dollar, from 22nd May 2013 to the end of 2013, be attributed to Federal Reserve policy?".

My investigation was undertaken by firstly creating a theoretical hypothesis as to what the answer of my research question would be. This was done using analysis of the supply and demand factors that could have affected the INR/USD exchange rate.

Secondly, I interviewed two experts, to obtain their viewpoints on the likely answer to my research question.

Thirdly, I undertook an empirical investigation of my research question. I looked for a 'control' economy, with its own currency, which I could compare to India. Comparing the performance of the two currencies allowed me to estimate what percentage of the rupee's decline could be attributed to the tapering announcement, and what percentage could be attributed to other causes.

My conclusion is that the depreciation of the rupee against the dollar from 22nd May 2013 to the end of 2013 can primarily be attributed to Federal Reserve policy. India's macroeconomic problems, worsening trade position with the U.S. and lack of prompt policy response to the rupee's weakness acted as secondary, but nonetheless significant, contributing factors.

Introduction

QE was implemented by the Federal Reserve in 2008 in a bid to increase the level of Aggregate Demand (AD), meaning total demand for goods and services, in the American economy. Under QE, the Federal Reserve undertook large-scale purchases of U.S. government bonds from commercial banks. Firstly, it was hoped that this would result in increased lending to firms for capital investment (the component of AD consisting of the purchase of man-made goods used to produce goods and services), since banks would have more cash available for this purpose. Secondly, QE was intended to drive up the prices of U.S. government bonds. Because each bond pays a fixed sum of money to its owner each year, this would reduce the 'yield' – the annual income paid to the owner of an asset divided by the asset's price – of U.S. government bonds, making them less attractive investments. The hope was that this would reduce the appeal to banks of putting their money into government bonds, increasing the relative attractiveness of lending to firms for capital investment. This can be thought of as an attempt by the Federal Reserve to induce a substitution effect, through increasing the cost of government bonds in order to increase demand for substitutes, specifically loans to firms for capital investment, so as to drive down long term interest rates and increase capital investment levels.

The hope was that the increased AD resulting from QE would boost U.S. economic growth. However, QE had adverse effects. It caused capital to flood out of the U.S. and into emerging markets such as India, where bonds and shares remained cheaper because governments had not implemented their own easing policies (unlike in most developed countries). This caused emerging market currencies, such as the rupee, to appreciate as demand for them increased.

On 22nd May 2013 then-chairman of the Federal Reserve, Ben Bernanke, announced that QE would soon be slowed down, or 'tapered', and stopped completely once the U.S. economy had returned to full capacity. (Market Watch, 2013) In anticipation of yields on U.S. government bonds rising as government demand for them decreased, capital began to flow out of emerging markets and back to the U.S. This increased demand for U.S. dollars in emerging market currencies such as the Indian rupee, causing the value of the dollar against these currencies to rise for the remainder of 2013. (XE 2014) In other words, these currencies depreciated against the dollar.

I will be considering how India's macroeconomic problems, specifically high inflation, slowing economic growth and a large government budget deficit may have exacerbated the rupee's depreciation against the dollar over this time period. These factors may have reduced the confidence of investors in the profitability of Indian firms, and consequently the value of Indian 'securities' (meaning tradable investments such as government bonds, corporate bonds, and shares in companies). They may also have reduced investor confidence in the ability of the government to service its debts. Decreased investor confidence in the value of Indian securities may therefore have contributed to outflows of capital from the economy, exacerbating the depreciation in the value of the nation's currency against the dollar.

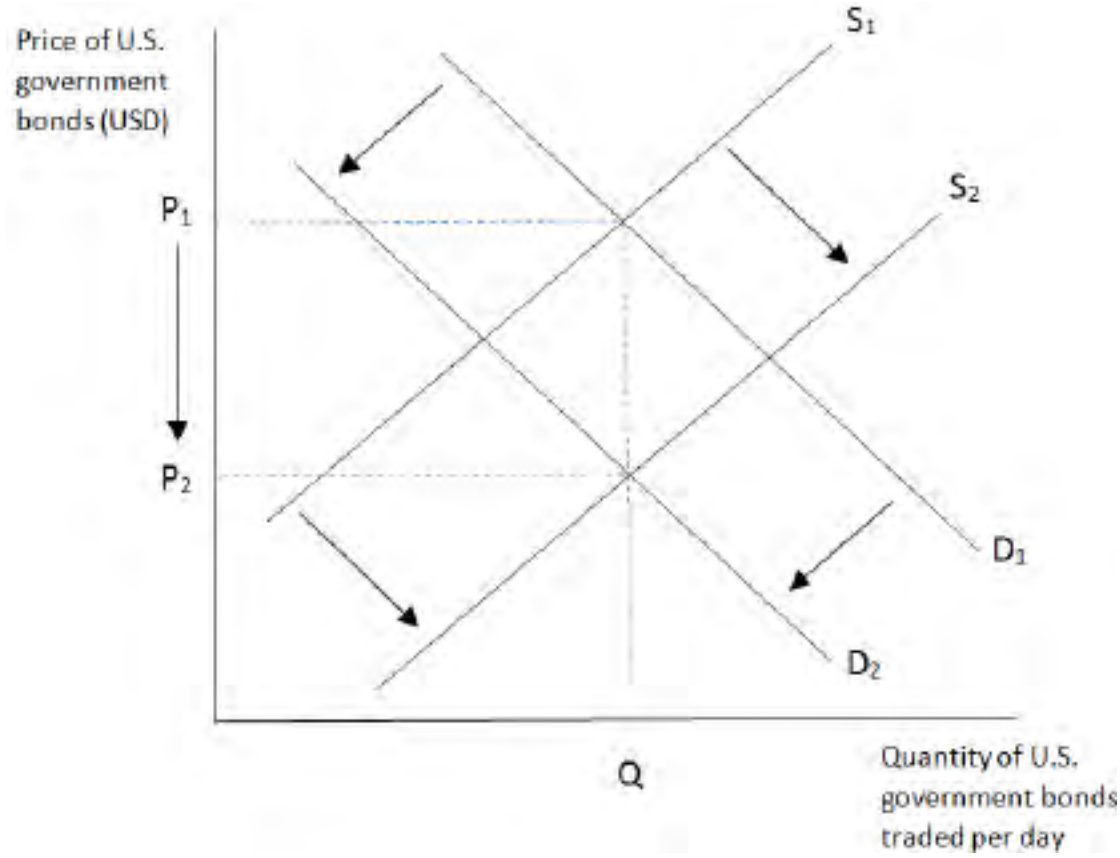
As well as considering the impact of speculative investor activity on the INR/USD exchange rate, I will also consider the impact of a worsening of India's trade position with the U.S. over the relevant time period. This worsening was driven in large part by India's high inflation rate, which made Indian exports to the U.S. less competitive and U.S. imports to India more competitive. This in theory would have increased demand for dollars from India and decreased demand for rupees from the U.S., contributing to the depreciation of the rupee against the dollar.

This research question is significant because it weighs into the ongoing debate regarding the extent of the impact of developed world central bank policy on developing economies. If the policies of central banks in developed countries are judged to have a strong destabilising effect on emerging economies, then a new system of international cooperation between central banks will likely be needed to ensure global macroeconomic stability.

Theoretical Hypothesis

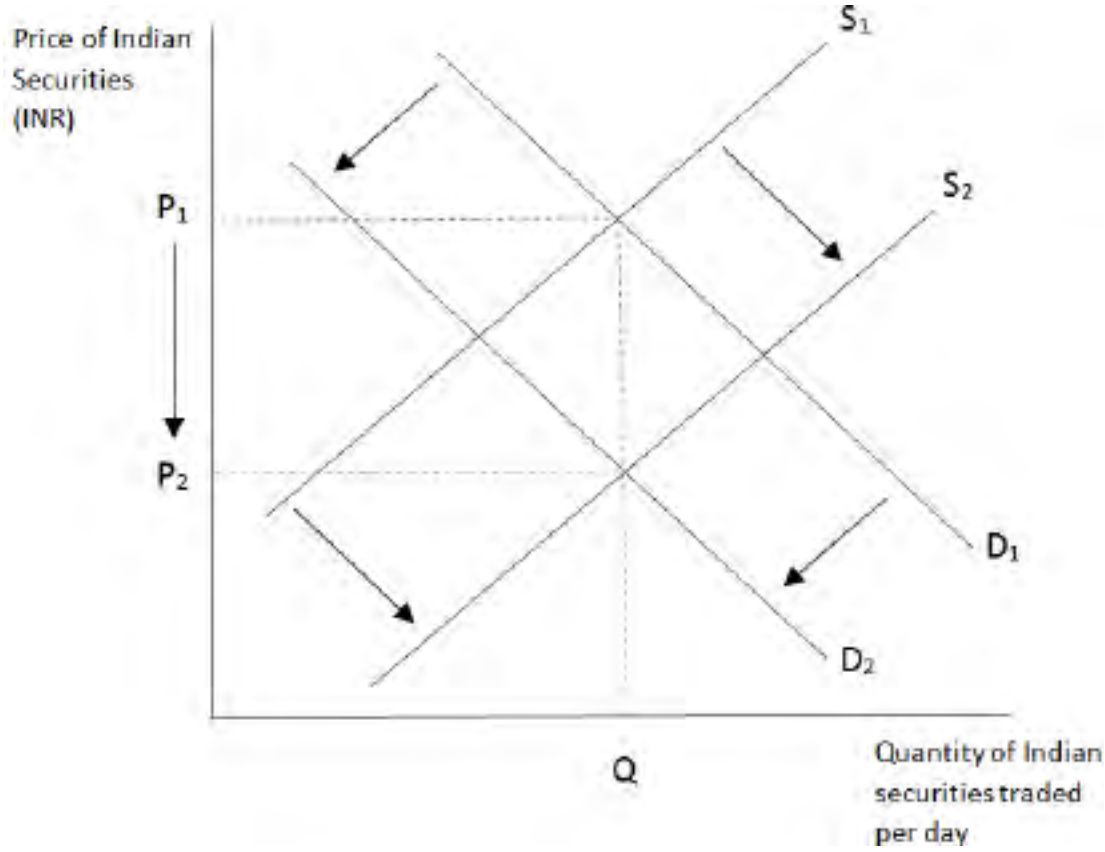
Hypothesis Regarding the Impact of Tapering

With the announcement that the Federal Reserve would soon begin tapering, I would expect that investors would anticipate the price falls that would result from tapering. By this, I mean that investors would realize that, ceteris paribus, demand for U.S. government bonds would soon fall and that the prices of these bonds would soon fall as a result of this. I would consequently expect to see many investors attempt to sell their U.S government bonds at the current price before prices fell. Investors would also become less willing to buy U.S. government bonds at their current price, as they would not want to purchase an asset that would soon decline in value. The effect of this is shown by the decrease in demand for U.S. government bonds and the increase in supply of them on the diagram below:

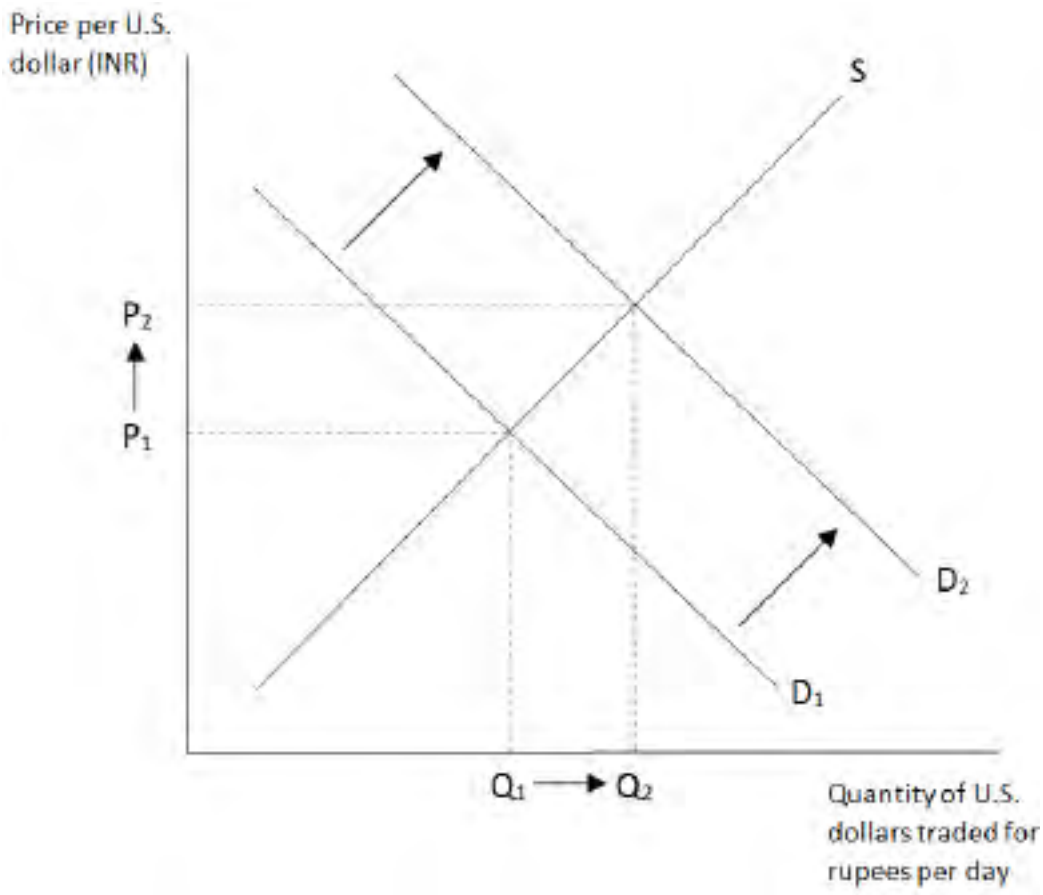


U.S. government bond prices would quickly adjust to the new equilibrium price, P_2, at which investors would no longer expect prices to fall any further as a result of tapering. Financiers would refer to this as the impact of the imminent tapering being 'priced in' to the bond market.

I would expect to see a fall in demand for substitutes for U.S. government bonds as U.S. government bond prices fell to P_2, as these substitutes would become relatively less attractive as investments in comparison to U.S. government bonds. More specifically, I would expect to see a decrease in demand for Indian securities as they would become relatively less desirable investments. I would also expect to see an increase in supply of these securities, as investors would attempt to sell them in order to move their capital into the U.S. instead. The effect of this on the market for Indian securities is shown below:



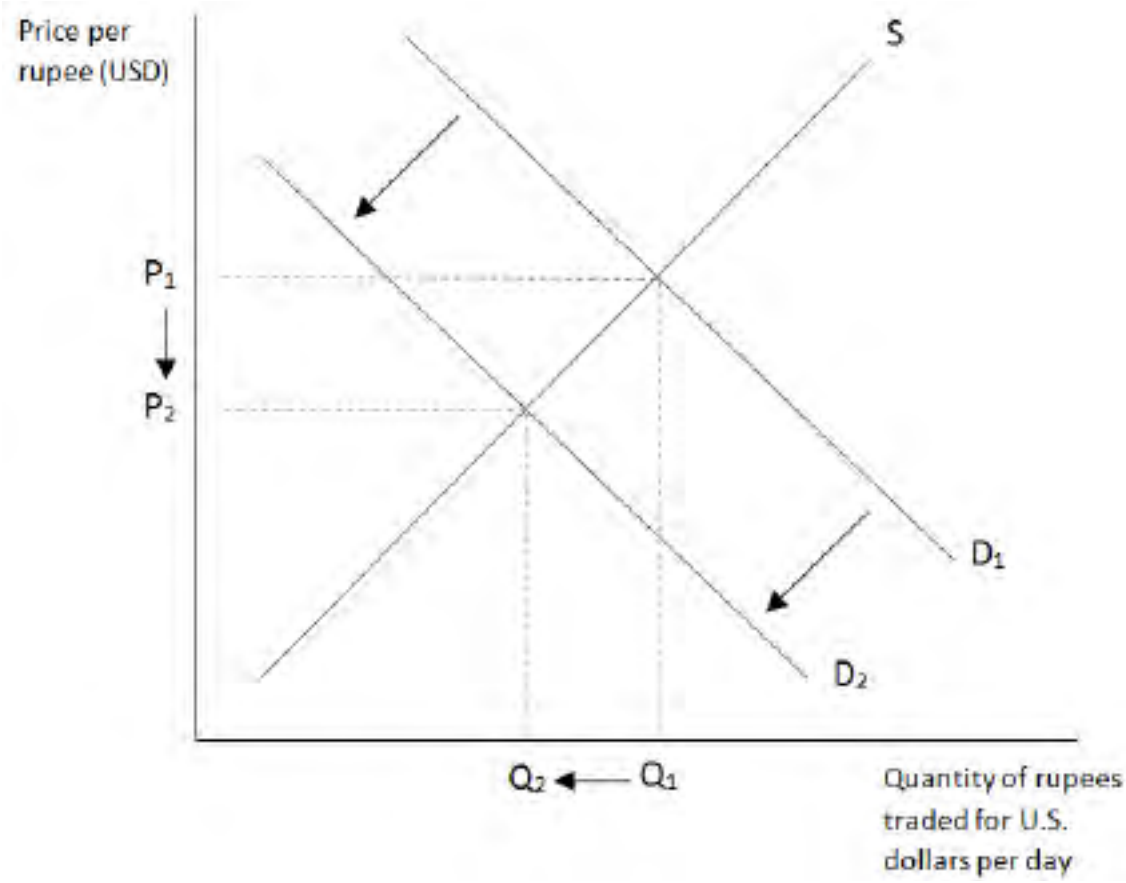
I would subsequently expect to see demand for U.S. dollars from investors in Indian securities to rise as these investors would need to exchange their rupees for dollars in order to purchase U.S. government bonds. The effect of this on the foreign exchange market is shown below:



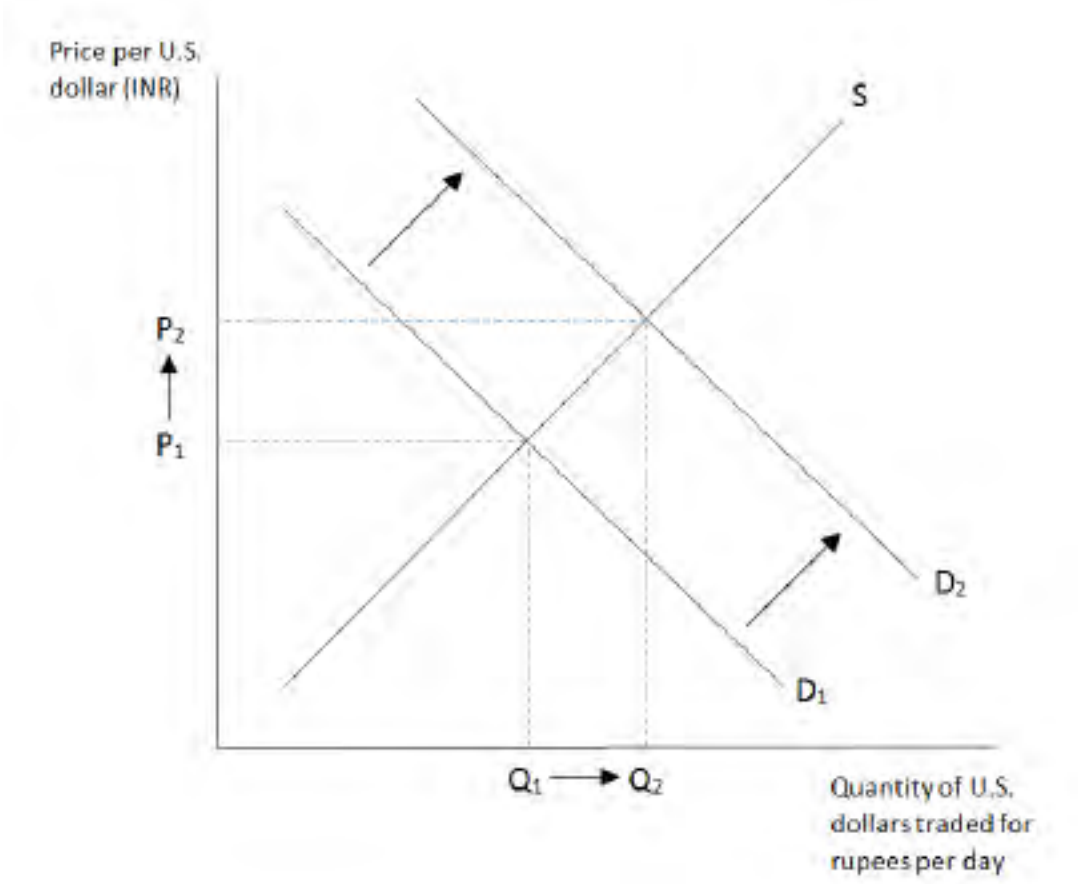
I would expect to see the price of each U.S. dollar to rise, in rupees, from P_1 to P_2; in other words, as a result of the tapering announcement, the rupee would depreciate against the dollar, meaning that each Indian rupee would be able to buy fewer U.S. dollars.

Hypothesis Regarding the Impact of India's Macroeconomic Problems

High inflation (meaning a rapid increase of the average price level) is one macroeconomic problem that may have contributed to the depreciation of the rupee against the dollar from May 2013 to December 2013, through its impact on India's trade balance with the U.S. (the difference between exports and imports to and from the U.S.). In 2013, India's consumer price index (CPI), a basket of goods commonly purchased by households that is used to estimate the average price level, rose by 9.13%, whereas the U.S. CPI rose by only 1.5%. (Inflation.eu 2014) Because the prices of goods produced in India were rising much faster than the prices of goods produced in the U.S., Indian exports to the U.S would have rapidly become less and less competitive in comparison to domestically produced American goods over the course of 2013. We would expect this to reduce demand for Indian exports. This would reduce demand for rupees from the U.S., decreasing the value of the rupee against the dollar as shown below:



Equally, high inflation in India compared to the U.S. would rapidly reduce the competitiveness of domestically produced goods for Indian consumers in comparison to imports from the U.S. I would expect this to increase demand for American imports, as they would become relatively cheap compared to domestically produced goods. This would cause an increase in demand for U.S. dollars from India, increasing the value of the dollar against the rupee as shown on the diagram below:



Slowing economic growth is another macroeconomic problem that may have contributed to the decline in the value of the rupee. In 2010, India's real GDP (the total output of the economy, adjusted for inflation) grew by a remarkable 10.09%. (Statista 2014) By 2012, however, growth had slowed to a less impressive 4.9%. (Statista 2014) The slowdown in economic growth may have damaged the confidence of investors for a number of reasons. Firstly, slowing economic growth could reduce the ability of the Indian government to service its debts, since lower output growth would mean slower growth in tax revenues, provided the level of taxation, as a percentage of real GDP, remained constant. Secondly, slowing economic growth would reduce growth in the profitability of firms, provided the share of GDP allocated to profits remained constant. This would reduce the ability of Indian corporations to service their debts, while also making shares in Indian companies less attractive. Reduced confidence amongst investors would likely result in many attempting to move their capital out of Indian financial markets and into U.S. financial markets, causing another rise in demand for dollars from holders of rupees and consequently contributing to the depreciation of the rupee against the dollar.

The problem of reduced investor confidence may have been worsened by the government's poor fiscal balance (meaning the difference between taxation and government spending). From 2012 to 2013 the Indian government ran a budget deficit (meaning government spending exceeding taxation) of 5.8% of GDP. (Trading Economics 2014) Because India is an emerging economy, and does not have the luxury of a long history of honoring its debts that can reassure investors, many would likely have become worried: they may have felt that the increase in government borrowing needed to finance this deficit would hurt the government's ability to service its existing debts. This may have made investors more likely to move their capital out of Indian government bonds and into less risky U.S. government bonds, again increasing demand for U.S. dollars from holders of rupees and causing the dollar to appreciate against the rupee.

Hypothesis Statement

I hypothesize that the depreciation of the rupee against the dollar from May 2013 to December 2013 can, to a considerable extent, be attributed to the policies of the Federal Reserve. However, I also hypothesize that India's macroeconomic problems and worsening trade balance with the U.S. were significant factors.

Research Methodology

Part One – Consultation with Experts

I will conduct interviews with two leading figures in the economic and financial worlds, in order to obtain secondary opinions on the answer to my research question. These will be used in conjunction with my primary research in order to formulate my conclusion.

Part Two – Confirmation of Theory

I will begin my primary research by investigating the main premises of my investigation. Firstly, I will determine the extent to which Ben Bernanke's tapering announcement led to capital outflows from emerging economies, since this will determine the extent of its impact on exchange rates. Secondly, I will determine the extent to which weak macroeconomic fundamentals may lead to capital outflows from emerging economies. Thirdly, I will determine the extent to which India's trade position with the U.S. deteriorated between May and December 2013.

Part Three – Statistical Comparison Between the Rupee and the Peso

Provided that the premises of my investigation hold true, I will compare the performance of the rupee against the dollar, from May 2013 to December 2013, to that of the Chilean peso. Chile, like India, is classed as an emerging market. Despite this, it suffers from none of the major macroeconomic problems that have bedeviled India: inflation is low, economic growth is stable, and the budget deficit is far smaller as a percentage of GDP than that of India. This means that it should not have suffered from capital flight driven by speculative investors as would have been expected in the case of India. Also, Chile's trade position with India did not deteriorate between May and December 2013. Finally, Chile has a reputation for political stability. I will therefore assume that depreciation of the peso against the dollar, from May 2013 to December 2013, was not significantly driven by trade or by speculative activity resulting from Chile's macroeconomic or political fundamentals. It follows from this that any depreciation of the peso against the dollar over the relevant time period must have been the result of Ben Bernanke's tapering announcement. By comparing the depreciations in the value of each currency against the dollar, I will obtain an estimate as to what percentage of the rupee's depreciation against the dollar can be attributed to tapering, and what percentage can be attributed to underlying macroeconomic problems.

Research Findings

Part One – Consultation with Experts

I interviewed Robert Raymond, a French macroeconomist, and David Tilles, a British investment manager. Both experts' Curriculum Vitae, as well as the transcripts of my respective interviews with them, can be found in the appendices.

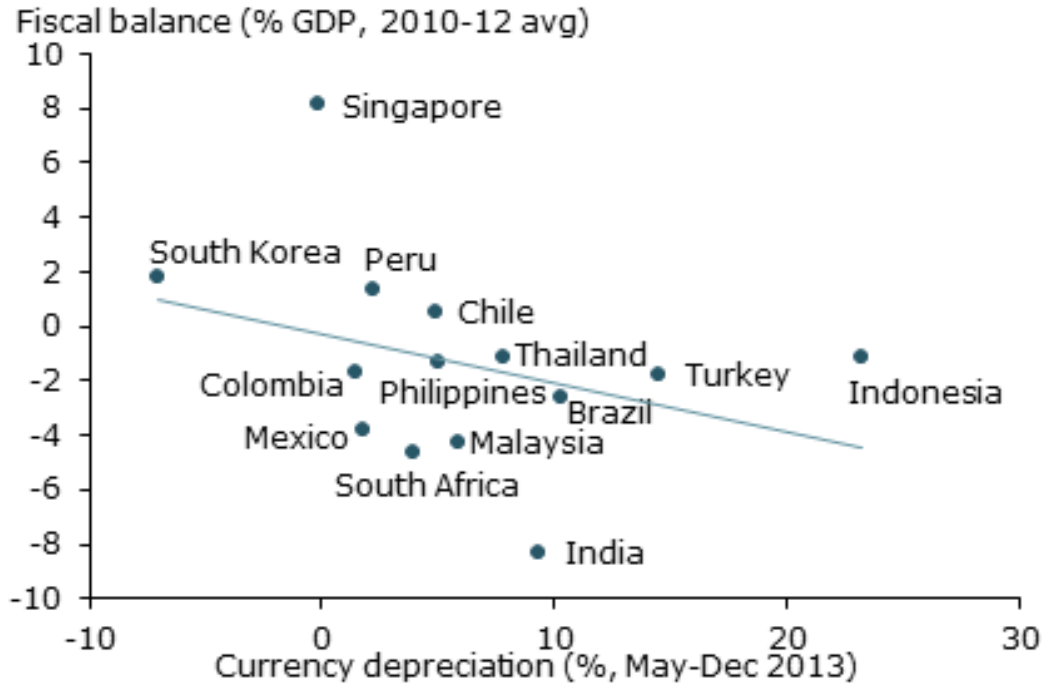
When questioned regarding their opinions of what caused the rupee's depreciation against the dollar, from 22nd May 2013 to the end of 2013, both interviewees initially talked at length about the tapering announcement, indicating that it was indeed the primary cause of the rupee's weakness. Both experts went on to mention the role India's macroeconomic problems played in the currency's weakness, but labeled these problems as factors that exacerbated the decline rather than being its root cause. Mr. Tilles also noted that the Reserve Bank of India's lack of a governor at the time, and its consequent lack of an appropriate response to the rupee's depreciation, served to weaken the currency further.

On the basis of these interviews, it would seem that Federal Reserve policy was the foremost cause of the depreciation of the rupee against the dollar from 22nd May 2013 to the end of 2013. India's macroeconomic problems and lack of prompt policy response to the rupee's weakness seem to have been secondary, though still significant, factors.

Part Two – Confirmation of Theory

Research indicates that Bernanke's statement regarding tapering did indeed result in large-scale sell offs of U.S. government bonds, as shown by the substantial fall in U.S. government bond prices and consequent rise in their yields that occurred immediately after the announcement was made (Powell, 2013). Rising yields on U.S. government bonds were soon followed by large capital outflows from emerging markets – over $25bn in June 2013 alone. In turn, all free-floating emerging market currencies, with the exception of the South Korean Won, experienced depreciations against the dollar over the next six months. (Nechio, 2014). This indicates that the main premise of my investigation – that Ben Bernanke's statement resulted in rising U.S. government bond yields, causing large capital outflows from emerging markets and depreciations in the value of their currencies – is correct.

Research also indicates that the second premise of my investigation – that weak macroeconomic fundamentals may contribute to capital flight from a country – also holds true. (Nechio, 2014) The graph below shows the relationship between an economy's fiscal balance and the performance of its currency in the seven months ending December 2013:



Source: Bloomberg

A negative correlation of -0.3 between a country's fiscal balance and the depreciation of its currency from May 2013 to December 2013 can be observed. (Nechio, 2014) This means that countries with large budget deficits generally saw their currencies perform poorly over this period, while countries with budget surpluses generally saw their currencies depreciate less. Admittedly, this does not confirm that the other macroeconomic problems faced by India, namely slowing economic growth and high inflation, contributed to the decline in the INR/USD exchange rate from May 2013 to December 2013. However, it does provide solid evidence that weak macroeconomic fundamentals may contribute to capital flight from emerging economies.

Finally, India's trade position did indeed worsen between May 2013 and December 2013. In the fiscal quarter ending August 2013, India's net exports to the U.S. totaled $5.32bn. In the next fiscal quarter ending December 2013, however, India's net exports to the U.S. totaled only $4.41bn - a reduction of $912m.

Capital flight resulting from the tapering announcement (and therefore Federal Reserve policy), capital flight resulting from India's weak macroeconomic fundamentals, and deterioration of India's trade position with the U.S., do therefore all seem to have contributed to the rupee's depreciation against the dollar from May 2013 to December 2013, as I predicted in my hypothesis. In the next section of my investigation, I will investigate the extent to which Federal Reserve policy alone was responsible.

Part Three – Statistical Comparison Between the Rupee and the Peso

In this section of my research, I will compare the performance of the rupee against the dollar, from May 2013 to December 2013, to that of the Chilean peso. As was explained in my research methodology I have chosen Chile because, like India, it is an emerging economy that experienced large capital inflows as a result of QE. However, it has none of the economic problems that India suffers from, and its trade position with the U.S. did not deteriorate from July to December 2014. It also has a reputation for political stability, as was mentioned by Robert Raymond during our discussion. Below is a table comparing the macroeconomic fundamentals of India and Chile, with U.S. figures also included for additional perspective:

COUNTRY	INDIA	CHILE	U.S.
Inflation (2013, as a %)	9.1%	3.0%	1.5%
Change in economic growth (2010-2013, as a %)	-5.6%	-1.4%	-0.6%
Budget Surplus (as a % of GDP, 2013)	-4.9%	1.0%	-4.1%
Change in trade balance with U.S (Q3-Q4 2013)	-$912m	$50.3	N.A.
Political Stability (Percentile global ranking)	12.32	60.19	65.88

Data provided by Inflation.eu, Trading Economics, World Bank, CIA and U.S. Census Bureau

The appendices provide graphs that should be consulted in order to contextualize these data points over the last ten years.

With regards to inflation, Chile's 3.0% rise in the average price level during 2013 pales in comparison to the 9.1% rise experienced by India. Equally, between 2010 and 2013 Indian economic growth slowed by a substantial 5.6%, a far greater slowdown than the 1.4% decrease in output growth experienced by Chile. In 2013 the Indian government ran a budget deficit of 4.9% of GDP. The Chilean government did not have a budget deficit, instead running a budget surplus of 1.0% of GDP. Its budget can therefore be said to have been sustainable - far more sustainable, in fact, than that of the U.S. government. This is remarkable, given that loans backed by the full faith of the U.S. government are considered to be 'risk free' assets. Finally, Chile is considered to be one of the world's more politically stable countries; its ranking in the World Bank's Worldwide Governance Indicator's Report places it in the top 40% of nations for political stability.

Given the overall picture of macroeconomic and political stability in Chile, I will assume that any significant capital flight from its economy was not driven by the nation's macroeconomic or political fundamentals, and was therefore a direct consequence of the tapering announcement.

Because Chile's trade position with the U.S. did not deteriorate in the latter half of 2013, any depreciation of its currency over this time period must have been the result of capital flows rather than trade deterioration. It follows from this, and my previous assumption, that any depreciation of the peso against the dollar in this time period was the result of capital flight caused by Ben Bernanke's tapering announcement. The table below shows the depreciations against the U.S. dollar experienced by the rupee and the peso from 20 May 2013 to 30 December 2013:

CURRENCY	RUPEE (INR)	PESO (CLP)
Closing exchange rate to USD 20th May 2013	0.01815	0.00206
Closing exchange rate to USD 30th Dec 2013	0.01617	0.00190
% Depreciation against the USD 23/05/13 to 30/12/13	10.9%	7.8%

Data provided by xe.com

This data implies that an economy with sound fundamentals and no deterioration in its trade position with the U.S. would have suffered a 7.8% depreciation against the dollar, from 20th May 2013 to 30th December 2013. However, we see that the rupee suffered a larger 10.9% depreciation against the dollar. Mathematically speaking, 71% of the rupee's decline can therefore be attributed to Ben Bernanke's announcement, with the remaining 29% being a consequence of capital flight resulting from India's macroeconomic problems and also the deterioration of its trade position with the U.S. This indicates that, for the most part, the rupee's decline in value against the dollar from May 2013 to December 2013 can indeed be attributed to Federal Reserve policy.

It must be acknowledged that this approach is not entirely rigorous. There are two problems with using Chile as a control economy. Firstly, Chile, while still being an emerging economy, has a relatively high real GDP per capita of $15,732, while India only has a real GDP per capita of $1,499. (World Bank 2013) This may mean that Chile is perceived as a less risky destination for investment, as was alluded to by one of my interviewees during my secondary investigation. In turn, this may have made its capital outflows lower and consequently its currency stronger, than would have been the case in an ideal control economy with the same GDP per capita as India. Secondly, Chile's trade position with the U.S. actually improved in the latter half of 2013. Since our ideal control economy would have seen no change in its trade position with the U.S., this again would have made the peso stronger than an ideal control currency. Attributing 71% of the rupee's decline to the tapering announcement is likely somewhat conservative in light of these points: Federal Reserve policy probably played an even more prominent role than my statistical analysis suggests.

Conclusion

Both my primary research and my interviews with experts indicate that the depreciation of the rupee against the dollar, from 22nd May 2013 to the end of 2013, can primarily be attributed to Federal Reserve policy. However, Federal Reserve policy was evidently not the only cause of the rupee's depreciation, with India's macroeconomic problems, worsening trade position with the U.S. and lack of prompt policy response to the rupee's weakness serving as secondary, but nonetheless significant, contributing factors.

My investigation leaves an obvious question unanswered; the relative importance of speculative capital flows driven by India's macroeconomic problems rather than the tapering announcement, and of India's deteriorating trade position with the U.S., in the rupee's depreciation against the dollar from 22nd May 2013 to the end of 2013.

Another question also arises from my findings. Because developed world monetary policy can be seen to have a destabilising impact on emerging economies, it seems that a system of closer cooperation between central banks is needed to ensure global macroeconomic stability. The question is how such a system of cooperation could be set up; it would clearly be a complex and difficult task, given the sheer number of central banks across the world. Interestingly, India recently announced that it would be setting up a supranational bank, with $100bn of capital to ease currency crises, alongside Brazil, Russia, China and South Africa. (DURDEN, 2014). This indicates that the developing world is indeed moving towards an era of closer cooperation between monetary authorities.

Bibliography

ANDERTON. (2006) *Economics Fourth Edition*

AGUILAR. (2013) *A short history of QE and the market*. Available from: http://projects.marketwatch.com/short-history-of-qe-and-the-market-timeline

BANYAN. (2013) *India's Economy: A five-star problem.* Available from: http://www.economist.com/blogs/banyan/2013/08/indias-economy?zid=309&ah=80dcf288b8561b012f603b9fd9577f0e

CIA. (2014) *The World Factbook.* Available from: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2222rank.html

DURDEN. (2014) *BRICS announce $100bn to bypass Fed, Developed World Central Banks.* Available from: http://www.zerohedge.com/news/2014-07-15/brics-announce-100-billion-reserve-bypass-fed-developed-world-central-banks

Inflation.eu. (2014) *Historic inflation – CPI inflation*. Available from: http://www.inflation.eu/inflation-rates/india/historic-inflation

NECHIO. (2013) *FRBSF Economic Letter*. Available from: http://www.frbsf.org/economic-research/publications/economic-letter/2014/march/federal-reserve-tapering-emerging-markets/

Statista. (2014) *India: Real gross domestic product (GDP) growth rate from 2004 to 2014.* Available from: http://www.statista.com/statistics/263617/gross-domestic-product-gdp-growth-rate-in-india/

Trading Economics. (2014) *Government Budgets.* Available from: http://www.tradingeconomics.com/india/government-budget http://www.tradingeconomics.com/chile/government-budget http://www.tradingeconomics.com/united-states/government-budget

United States Census Bureau. (2014) *U.S. Trade in Goods by Country*. Available from: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2222rank.html

World Bank. (2013) *Growth in India.* Available from: http://web.worldbank.org/WBSITE/EXTERNAL/COUNTRIES/SOUTHASIAEXT/EXTSARREGTOPMACECOGRO/0,,contentMDK:20592481~pagePK:34004173~piPK:34003707~theSitePK:579398,00.html

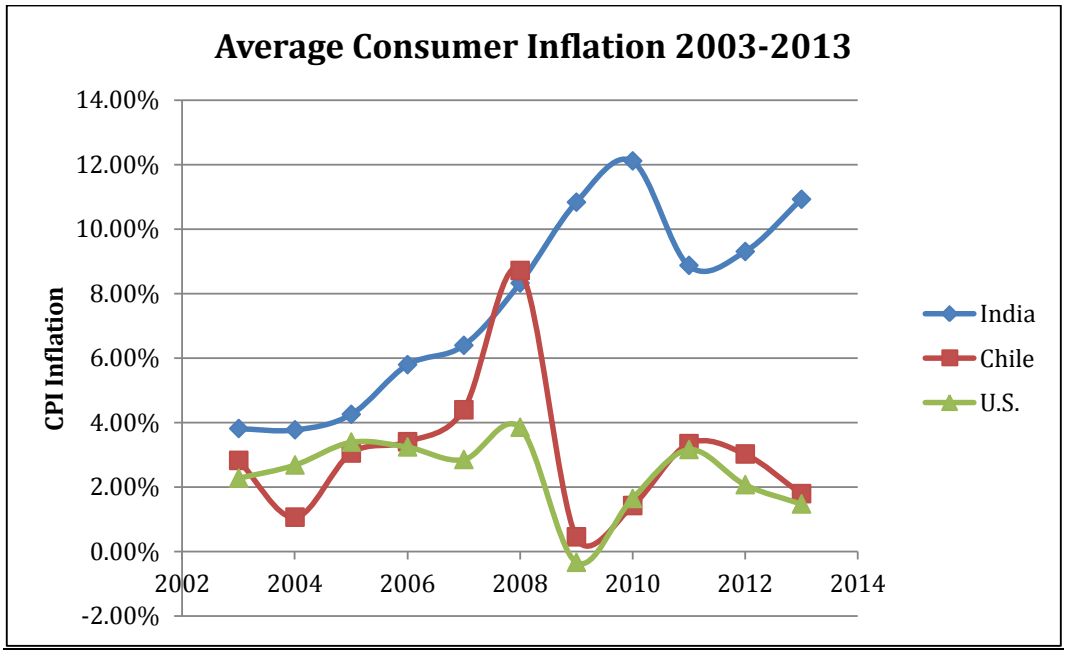
World Bank. (2013) *Worldwide Governance Indicators.* Available from: http://info.worldbank.org/governance/wgi/

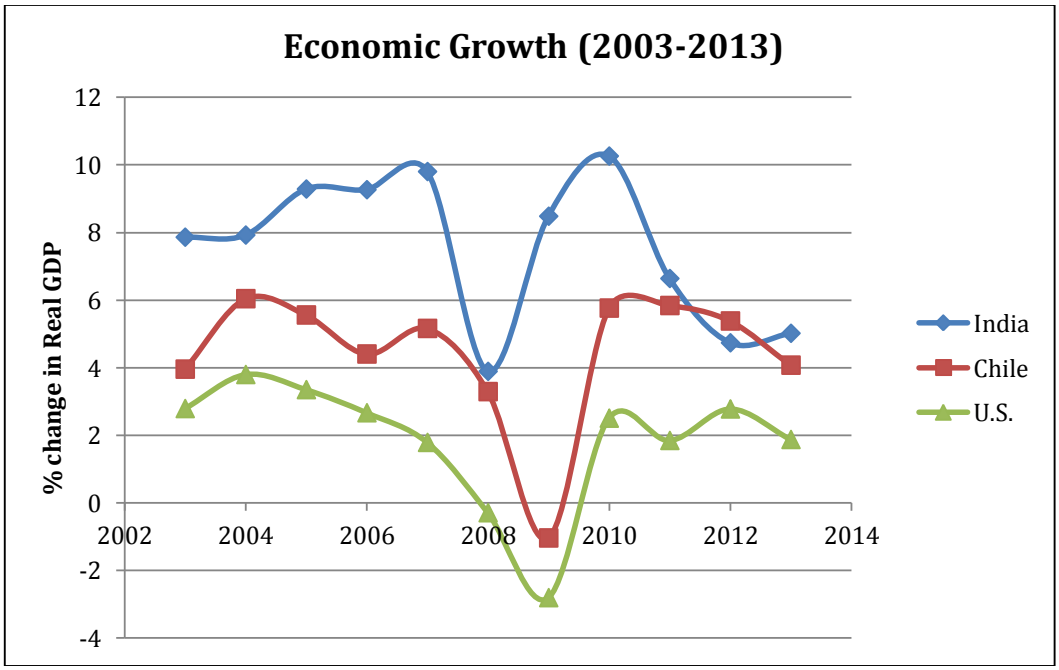
Xe.com. (2014) *Current and historical rate tables.* Available from: http://www.xe.com/currencytables/

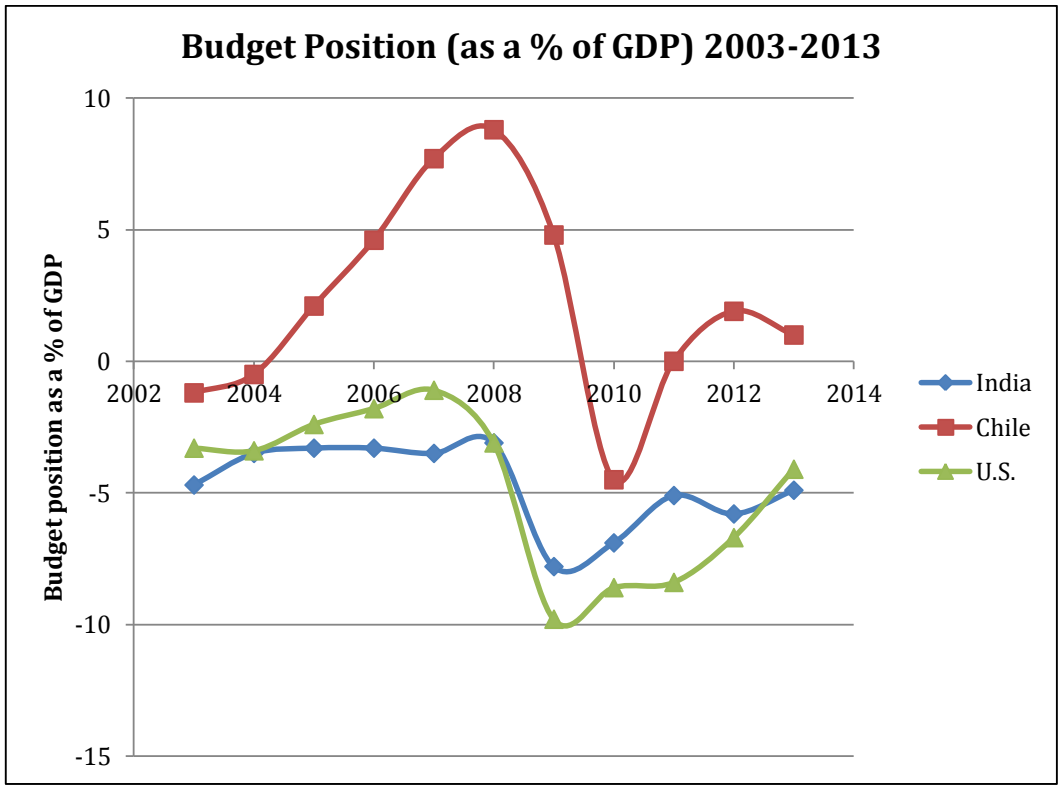
Appendices

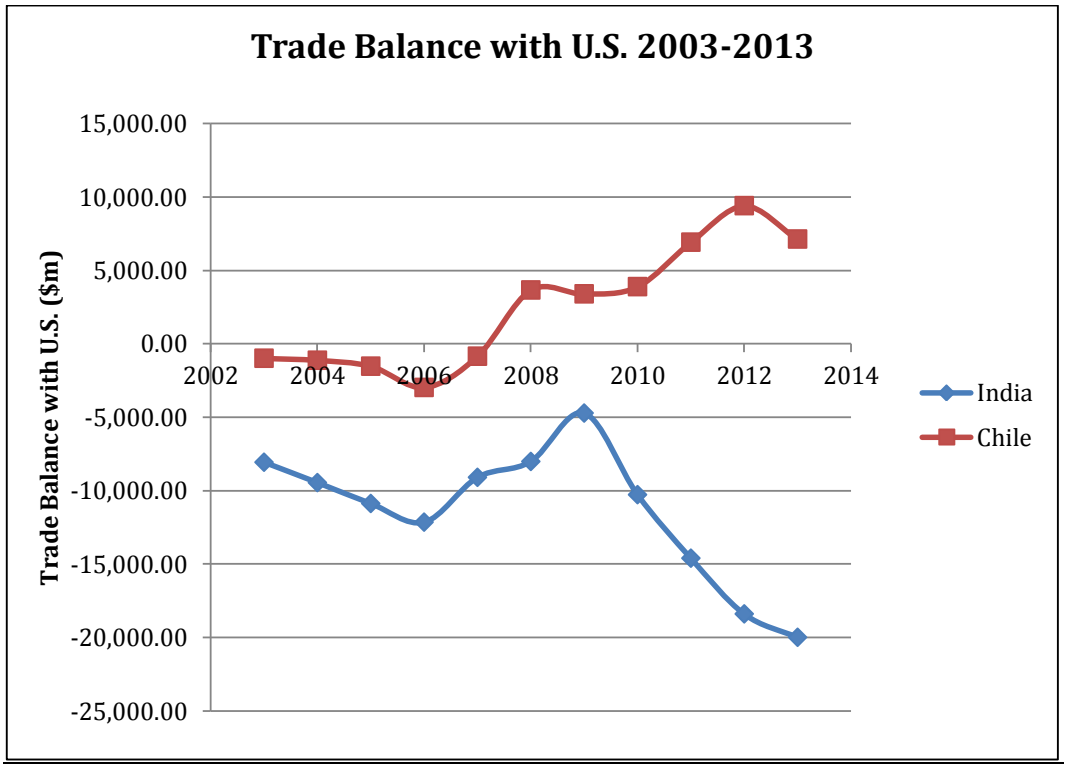
Relevant Macroeconomic Data

The following graphs show macroeconomic data for India, Chile and the U.S. This data can be used to contextualize the data points referred to in the body of my essay. Data was provided by inflation.eu, the World Bank, Trading Economics and the U.S. Census Bureau.






Budget Position (as a % of GDP) 2003-2013
Budget position as a % of GDP
10
5
0
-5
-10
-15
2002
2004
2006
2008
2010
2012
2014
India
Chile
U.S.


Trade Balance with U.S. 2003-2013
Trade Balance with U.S. ($m)
15,000.00
10,000.00
5,000.00
0.00
-5,000.00
-10,000.00
-15,000.00
-20,000.00
-25,000.00
2002
2004
2006
2008
2010
2012
2014
India
Chile

Robert Raymond Curriculum Vitae and Interview Transcript

Mr. Raymond is a French expert in macroeconomics, having served as Director General at the European Monetary Institute, Senior Special Advisor to the Executive Board of the European Central Bank (ECB), and the ECB's permanent representative in Washington D.C. with observer status at the International Monetary Fund. Notably, Mr. Raymond co-wrote an article examining the consequences of central bank independence, which included analysis of the impacts of developed world monetary policy on emerging economies.

MK: What is your opinion of India's macroeconomic fundamentals?

RR: India is a land of contrast.

It can draw on a number of comparative advantages due to a vast territory with significant resources (although little energy), an abundant and cheap manpower as well as a long lasting tradition of entrepreneurship, however so far limited to some segments of the economy, and finally a democratic political system. Like other large Asian economies, India recently enjoyed several years of high growth, in particular during the aftermath of the 2008 crisis in the USA and Europe. Nevertheless it is not quite on the first row. Indeed it could do better if it were able to overcome various structural obstacles.

One is the difficulty in training this large and ever growing population. Besides some remarkable achievements, for instance in the IT sector, a large part of the inhabitants keep a low level of education. Rural communities remain archaic and cultural standards are resilient. Infrastructures are deficient. The inefficiency of the administration and the large competence of the local governments within this Federal State are an obstacle to a rapid modernization, in particular in the transportation sector and basic industries. It will take time to reach a consensus to change these features.

In a nutshell: India has a large potential which is not yet fully exploited.

MK: Please could you give your opinion of Chile's economy?

RR: Although it is included into the category of developing countries, Chile has some common features with European countries. It is a middle-size nation, with a small population (16 millions) which is about half of a European ancestry and is aging, contrary to what can be observed in most emerging economies. It has become a democracy after a long dictatorship. The mental of the inhabitants seems to be less volatile than in some Latin American neighbor countries (in Europe, Germany also, which enjoys a reputation of stability, had a terrible dictatorship in the 20th century).

The economy is reasonably solid, enjoying strong resources of copper and a well-developed agriculture. Inflation remains manageable, growth is good (around 6%) and the rate of unemployment is at 8%, a level that a few European economies could envy.

These fundamentals are a good basis for the future, depending of the quality of the political management in the years to come, the ability to avoid "el mal gobierno".

MK: What do you think caused the depreciation of the rupee against the dollar from May 2013 to December 2013?

RR: The financial and economic crisis which started in 2008 in the most developed economies led their central banks to increase liquidity and bring interest rates close to 0. Therefore investors of this part of the world, looking for a better return, massively invested money in the developing countries which enjoyed a high rate of growth and a large foreign trade surplus. They revised drastically their opinion in 2013, when it became likely that growth (at least some growth) would resume in the USA and lead the Federal Reserve System to taper its Quantitative Easing policy, leading to higher interest rates, as foreseen by its chairman Ben Bernanke on May 22. This triggered a selloff of the portfolios of financial assets issued in developing countries. The outflow of capital was more acute in India than in other underdeveloped economies, because of the distrust of investors inspired by the poor current account position, put in deficit by the macroeconomic policies supporting consumption and the rural areas previously implemented by the Indian Federal Government. Financial markets calmed down when the Fed clarified its policy on 30 September 2013, announcing a cautious approach in reversing its policy stance.

David Tilles Curriculum Vitae and Interview Transcript

David Tilles is the founder and Executive Chairman of Mondrian Investment Partners. Mondrian currently actively manages $75bn for clients, with over $18bn invested in emerging markets including India and Chile.

MK: In terms of macroeconomic fundamentals, what is Mondrian's opinion of India as a destination for investment?

DT: India has the potential for high and sustained economic growth given its low level of development, low urbanisation and good demographics. However one must balance against this the persistently high inflation, sustained fiscal deficit and current account deficit which create risks for foreign investors, particularly from the risk of currency depreciation. India has also suffered from unstable governments and poor governance systems which creates additional risk for investors. India is a high-risk/ high return proposition for investors.

MK: In terms of macroeconomic fundamentals, what is Mondrian's opinion of Chile and a destination for investment?

DT: Chile is a more stable and predictable destination for investment. The country is already relatively highly developed so the outlook for economic growth is less exciting. Chile is fiscally responsible and has a strong respect for institutions and adherence to rules, so many macroeconomic risks within the control of the country are carefully managed. But not all risks are within their control. As a small country with an open economy, Chile is highly dependent on imported goods which creates an imbalance in external accounts. The currency is particularly vulnerable if the price of copper falls while the price of oil rises since this can stoke the current account deficit. On balance Chile is a reasonably safe destination for investment but unlikely to deliver high returns.

MK: Has tapering of QE by the Federal Reserve impacted Mondrian's investment strategy in geographical terms?

DT: Yes - we have reduced exposure to markets that are negatively impacted by withdrawal of USD liquidity and higher rates, mainly those with large current account deficits like India, Indonesia, Turkey and Brazil. We also increased exposure to more stable countries like Malaysia and Taiwan.

MK: Did Mondrian choose to reduce its exposure to Indian investments from May 2013 to December 2013? If so, what was the main reason for this decision?

DT: Net-net we were slightly less exposed to India in December compared to May, having increased exposure during the summer months when the market was very weak and then trimming through the autumn as it recovered. Across our portfolio aggregate exposure to Indonesia, India and Turkey proved to be more volatile and also more highly correlated than we expected, hence carrying higher risk than we had previously thought. We therefore reduced exposure as the Indian market recovered.

MK: What do you think was the primary cause of the depreciation of the rupee against the dollar in the latter half of 2013?

DT: The main cause of the depreciation of the rupee in the latter half of 2013 was capital flight. As the US moved to taper QE the interest rate differential was expected to come down significantly and the vulnerability of the rupee was exposed. This happened at a time when the central bank had no governor (hence no move to prop up the rupee) and market liquidity was thin, which magnified the movement in the currency and created a sense of panic. Some of the capital returned later in the year as market conditions improved and the central bank took corrective policy action.

Are Chinese Medicines safe? An investigation of the amount of heavy metals present in Chinese patent herbal medicines.

Laura Lau - Chemistry

The Extended Essay, in which I examined the safety of alternative medicines with the use of analytical chemistry, has exposed me to wider knowledge in Chemistry. From planning, designing the method, conducting experiments, to processes of analysis and evaluation, the investigation required patience, time and diligence. In completing this piece of research that included a number of failed experiments and frustrating errors, I believe having a genuine interest in the subject was instrumental. I found it very enjoyable being able to spend time researching and experimenting on the topic that I like - I even had the chance to visit the University of Greenwich to use their optical emission spectrometer! Although the Extended Essay only offers a brief glimpse of what research actually is, it consolidated my understanding and passion towards the subject area that I consider to pursue in the future, and it was ultimately very rewarding having completed my first research project.

Supervisor: Dr Rachel Yu

One of the oldest forms of medical practice, traditional Chinese medicines have existed for over 2000 years and have recently become more popular in the West. However, there are growing concerns about the potential toxicity of these products. For her essay, Laura came up with the ambitious idea of measuring the amount of toxic metals in some of the popular Chinese medicines. After thorough literature research, she carried out preliminary trials to set up parameters for her experiments and was able to perform experiments well beyond school level in both accuracy and efficiency. The calibre and precision of her research, however, quickly outstripped the ability of our own school-level equipment in trying to measure the level of toxins in her samples. Laura remained positive and showed real initiative whilst on a School Analyst Competition at the University of Greenwich, she discussed her own research with scientists in the Chemistry Department. Impressed, they then taught her to use their inductively coupled plasma optical emission spectrometry equipment, allowing her to develop more precise measurements. Throughout the project, Laura showed great enthusiasm and dedication to hard work. She has the curiosity, independence and ability to reflect on her own thinking which will enable her to become an excellent scientist in the future.

Contents

Abstract

Objective: This essay sets out to assess the safety of some Chinese premixed patent herbal pills, by identifying and quantifying the amount of toxic metals present, using the research question: *"Are Chinese Medicines safe? An investigation of the amount of heavy metals present in Chinese patent herbal medicines."* **Method:** The amounts of toxic metals in five herbal pills are examined by (1) precipitation of mercury as sulphide (for mercury specifically) and (2) inductively coupled plasma optical emission spectroscopy, which requires prior digestion by either solely heated with HNO3 or microwaved. **Results:** Results are compared with various limits set by food regulation organisations: FAO/WHO Expert Committee on Food Additives, European Medicines Agency, and Chinese Medicine Council of Hong Kong. Cobalt levels are below detection limit, while mercury levels are high in all pills and exceed tolerable intake levels. Content of arsenic and cadmium in pills 1 and 2 also exceed tolerable intake regulations. Other metals are present in lower levels and do not exceed any limits. Results of digestion by heating only and microwave are compared, with microwave digestion containing higher levels of heavy metal. Samples 1 and 2 on the whole, contain highest amounts of heavy metals. The results of mercury precipitation cannot be quantified due to the small amount formed, along with contamination and incomplete reaction. **Conclusion:** Heavy metals examined in ICP-OES exceed tolerable intake levels and reflects the need of stricter production quality control of herbal medicines in Asia. However it is unclear if the intensity of metals in the pills would result in more complex reactions. More investigations in the field of Traditional Chinese Medicines are required as well.

Introduction

It was when my grandmother had a stroke that I considered investigating the toxicity of some Chinese herbal medicines in my extended essay. As a traditional Chinese family, relatives recommended that we give her 'Angong-niuhuang pills' (herbal pills containing cattle gallstones) which claim to be very effective in 'detoxing' a patient suffering from stroke. However, we were aware of the reported toxicity/ contamination in these products – some herbal medicines are said to contain high amounts of mercury – therefore we were deterred from using these pills (Anon., 2013). This lead me to give thought to the compositional purity of the Chinese herbal pills that are widely available in Asian markets. This became an even more interesting question when I considered the growing popularity of Chinese herbology worldwide, and therefore enabled me to produce the following research topic on analytical chemistry for this investigation:

> *Are Chinese Medicines safe? An investigation of the amount of heavy metals present in Chinese patent herbal medicines.*

Traditional Chinese medicine incorporates a broad range of medicinal practices more than 2000 years old which, from the modern scientific perspective, are considered pseudoscience. Herbal medicine, the focus of this investigation, is only one aspect of it. The public generally believes that herbal medicines are milder and benign alternatives to Western drugs, as they are made from natural ingredients.

The five herbal pills examined in this investigation are common premixed patent medicines that are readily available in supermarkets, pharmacies and healthcare stores across Asia. The main uses are to "clear away heat", "detoxify", "relieve convolution" or "resuscitate". They are said to contain significant amounts of arsenic and mercury seen from atomic absorption spectroscopy (Espinoza, et al. 1996). Three of the five pills contain cinnabaris and realgar, which are suspected sources of arsenic and mercury. Owing to the fact that quality control policies in China are not strictly carried out, it is also possible that the improper manufacturing procedures could lead to further heavy metal contamination of the medicines (WHO, 2005). The heavy metals examined in this investigation are arsenic, cadmium, cobalt, chromium, copper, iron, mercury, molybdenum, nickel, lead and zinc. These heavy metals are potentially harmful to the human body as they interfere with metabolic processes, and they could also bioaccumulate in the body, cause neurological damage or even induce cancer.

The presence of such metals is measured by inductively coupled plasma optical emission spectrometry (ICP-OES) and, mercury specifically, will be precipitated as sulphide by bubbling hydrogen sulphide through dissolved samples.

Materials and Methods

The Herbal Pills

The following five pills are very conveniently purchased in Mannings healthcare store in Hong Kong. Table 1 shows their names, manufacturer, place of manufacture and uses. Wrapped in gold leaf, they are roughly spheres weighted 3 g, which contain plants, animal components and minerals mixed together manually. The pills are consumed by dissolving in warm water or wine. All medicine labels that come with the pills advise not to take them during pregnancy, and recommend doctors' consultation.

Table 1

Details of Herbal Pills from Medicine Labels

Samples	Name	Manufacturer	Place of manufacture	Uses	Directions	Composition
1 (Anon. 2009)	Angong-niuhuang pills (ox bezoar calming pills)	Beijing Tong Ren Tang	Hong Kong	Clear away the heat, detoxify, relieve convulsion and resuscitate. Effective for cerebro-vascular diseases, encephalitis, pneumonia and dysentery.	Taken orally, 1 pill each time, 1 time a day; 1/4 pill for children under 3 years old; 1/2 pill for children 4-6 years old. Don't use it for over 3 days or follow the physician's advice.	Calculus Bovis, Moschus, Margarita, Cinnabaris, Realgar, Rhizoma Coptidis, Radix Scutellariae, Fructus Gardeniae, Borneolum Syntheticum, Radix Curcumae
2 (Anon. 2009)			Beijing			
3 (Anon. 2007)	Angong-niuhuang pills (ox bezoar calming pills)	Hong Kong Ma Pak Leung	Hong Kong	Remedy for 'heatevil' fever and phlegm heat obstruction.	To be taken by mouth, or dissolve pill in warm boiled water. Adult: 1 pill 2 times a day, for severe condition, take 1 pill 3 times a day. Children: take half of adult dose. Not to be taken during pregnancy. Avoid cold, raw or greasy food while taking the pill. Stop when having a fever or flu. If condition is not improved, consult a medical doctor.	Arisaema Cum Bile, Rhizoma Coptidis, Bombyx Batryticatus, Radix Saposhnikoviae, Borneolum Syntheticum, Bovis Calculus, Moschus, Scorpio, Mel, Rhizoma Pinelliae Praeparatum, Radix Trichosanthis, Succinum, Radix Scutellariae, Fructus Amomi
4 (Anon. 2009)	Dahuoluo pills (Vessels reviving pills)	Beijing Tong Ren Tang	Beijing	Expelling wind and cold in human body, eliminating dampness. Effective for hyperlipidemia cerebrovascular accident or hemiparalysis, and rheumatism, chronic rheumatoid arthritis, rheumatoid arthritis, degenerative arthritis.	To take orally with warm water or yellow wine, 1-2 pills a time, 2 times a day; follow the physician's advice.	Agkistrodon (processing with wine), Zaocys (processing with wine), Radix et Rhizoma Rhei, Rhizoma Gastrodiae, Rhizoma Arisaematis (processed), Radix Rehmanniae Preparata, Sanguis Draconis, Radix Aconiti Kusnezoffii Preparata, Moschus, Calculus Bovis, bones of leopard (processed), Cornu Bubali concentrated powder.
5 (Anon. 2009)	Niuhuang-qingxin pills (ox bezoar soothing pills)	Beijing Tong Ren Tang	Beijing	Replenish the vital energy and the blood, relieve convulsions, tranquilize nervous strain, eliminate sputum and endogenous wind-syndrome.	Orally taken, 1 pill each time, 1 time a day; 1/4 pill a time for 3 years old children; 1/2 pill a time for 4-6 years old children; Follow the guidance of doctor.	Calculus Bovis, Cornu Saigae Tataricae, Moschus, Radix Ginseng, Rhizoma Atractylodis Macrocephalae (stir-frying with bran) Radix Angelicae Sinensis, Radix Paeoniae Alba, Radix Bupleuri, Rhizoma Zingiberis, Colla Corii Asini, Radix Platycodonis, Cornu Bubalipowder

Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES)

Background Information

ICP-OES (Perkin Elma Optima 4300 DV) set with operating conditions and parameters shown in table 2, was used to examine the digested samples. This analytical technique uses argon plasma, in which electric currents produced by electromagnetic induction supplies energy to dissociate digested samples and emit photons.

A Tesla unit first ionizes argon gas, which then flows towards an intense electromagnetic field generated by the radio frequency generator surrounding the ICP torch, to create Argon plasma (Zand, 2011).

Before being injected into the spectrometer directly, solid samples require acid digestion and are converted into wet aerosol.

The samples are then pumped into the analytical nebulizer in the spectrometer, and introduced into the plasma as mist, where it collides with the charged carriers of the plasma. A calibration blank containing internal standards are injected into the spectrometer and analysed before other samples. This allows the calibration of metal intensity in the digested samples (O'Haver, 2014). The constituent atoms or ions of the samples are then excited to a higher energy level, and emit photons of a characteristic wavelength, which are detected and recorded by a photomultiplier tube. The certain intensity of photons released reflects proportionally the concentration of elements present (Royal Philips, 2013).

Table 2

ICP-OES Operating Conditions	
Power	1300W
View mode	Axial
View distance	15mm
Read delay time	60sec
Replicates	3
Plasma gas flow	15L/min
Auxiliary gas flow	0.2L/min
Source equilibrium delay	15sec
Nebulizer	0.7L/min
Sample flow rate	1.5mL/min
Sample wash rate	1.5mL/min
Sample wash time	30sec
Background correction	2-point
Analytical Lines Wavelength	
As	188.979
Cd	228.802
Co	228.616
Cr	267.716
Cu	327.393
Fe	238.204
Hg	253.652
Mo	202.031
Ni	231.604
Pb	220.353
Zn	206.200

Procedural Write-up

Materials

- Large bucket for soaking glassware
- 7 boiling tubes
- 7 50ml glass beaker
- 7 watch glasses
- Chemicals:
    - Concentrated nitric acid (70%)
    - Concentrated hydrochloric acid
    - Medicine samples
    - Deionized water

Sample Preparation

Wet Digestion with HNO_3

Soak glassware in a pot of 10% nitric acid overnight to reduce contamination from glassware. Weigh out accurately 0.5 g of sample and add 10 ml of concentrated nitric acid (70%) to a 50 ml glass beaker. The higher the grade of nitric acid, the more accurate the results will be (D. Wray, pers. comm. 10th June 2014). Place the beaker on a hotplate in a fume cupboard, cover the beaker with a watch glass and heat for 3 hours at 70 °C. Allow to cool and add 20 ml of deionized water to the beaker, then filter into boiling tubes with a funnel and filter paper. Include two blanks containing 70% nitric acid and 20 ml deionized water.

Microwave Digestion with HNO_3

Weigh out 0.5g samples accurately into pressurized vessel (XP-1500 Plus). Add 10ml 70% nitric acid into the vessel and allow samples to predigest by standing open for a minimum of 15 minutes before sealing vessels. Seal and insert vessels into microwave oven (CEM MARS 5®), set microwave to digestion conditions shown in table 3, to enable complete digestion. Allow 15 minutes for cool down. Include two blanks, of 10ml 70% nitric acid. (Wang, et al. 1999).

Table 3

Microwave Heating Program Digestion Conditions	
Stage	1
Power	400W - 100%
Ramp time	15min
Pressure	800psi
Temperature	200°C
Stir	Off
Hold time	15min

A sample of all five pills and two controls was wet-digested with HNO_3 (samples 1-5, C_1, C_2), and another sample of pills 1, 5 and two other controls were digested by microwave (samples 1x, 5x, Cx_1, Cx_2). This is because pills 1 and 5, as suggested in literature (Espinoza, et al. 1996), has highest amounts of mercury and arsenic. Each sample is analysed three times, and the spectrometer is rinsed with 10% HNO_3 between every sample detection.

Determination of mercury as sulphide

Pills 1, 2 and 3 are examined by this method, as they are similar in content and function.

Background Information

Mercury (Hg) can be determined as mercuric sulphide (HgS):

$$Hg + H_2S \rightarrow HgS + H_2$$

Procedural Write-up

H_2S Production

Hydrogen sulphide is produced by reacting iron sulphide with hydrochloric acid:

$$FeS + 2HCl \rightarrow H_2S + FeCl_2$$

Materials for H_2S production

- Mortar and pestle
- Bung and connecting tube
- Electronic balance
- 1 Conical flask
- 1 200 ml beaker
- 1 Crucible
- Chemicals:
    - Hydrochloric acid (2M)
    - Iron(II) sulphide

Iron sulphide pellets are first crushed to powder form using pestle and mortar to increase the surface area for reaction. In a fume cupboard, react roughly 15 g of FeS powder with 150 ml HCl (1M) in a conical flask and bubble excess H_2S into sample immediately through connecting tubing, swirling the flask throughout.

Precipitating mercury

Materials

- Bungs and connecting tubes
- Electronic balance
- Magnetic stirrer and heat plate
- 6 Boiling tubes
- 3 100 ml beakers
- 2 50 ml measuring cylinder
- Grade 4 filter paper
- 3 Funnels
- Crucibles
- Boiling tube rack
- Stopwatch
- Chemicals:
    - Hydrochloric acid (1M)
    - Distilled water
    - Hydrogen sulphide
    - Medicines samples

Dissolve 0.5g of samples in 10 ml HCl (2M) in a 100 ml beaker, and then add 40 ml distilled water. Heat and stir using magnetic stirrer for 20 minutes. When cooled, filter out the insoluble components with funnel and filter paper, then separate equally into two boiling tubes, marking them as 'A' and 'B' and pass H_2S into 'A' boiling tubes using connecting tubing in the fume cupboard for 20 minuets. Keep the bung on the boiling tube loose. When finished, filter out the precipitation of all samples with funnel and grade 4 filter paper. Compare precipitation on filter paper of the same solution with (A) and without (B) H_2S bubbled through on filter paper, or measure the mass of precipitation on filter paper.

Results and Data Interpretation

ICP-OES

Results

Table 4 lists the mean intensity of wavelengths detected from the three repeats in mg/L. Table 5 shows the wavelengths intensity of each metal in mg/L, after subtracting the background noise in detection represented by the controls: the mean of C_1 and C_2 for samples 1-5; Cx_1 and Cx_2 for 1x and 5x. Table 6 shows the metal content in each sample in mg/kg, allowing the comparison of metal content in each sample.

Table 4

Mean Corrected Intensity of Wavelengths (mg/L)											
Samples	As	Cd	Co	Cr	Cu	Fe	Hg	Mo	Ni	Pb	Zn
1	516.5	17.49	0.022	0.126	0.428	2.539	49.26	0.028	0.073	0.641	1.829
2	488.8	16.33	0.024	0.159	0.266	2.636	40.98	0.028	0.097	0.109	1.167
3	1.302	0.033	0.024	0.197	0.863	1.619	7.322	0.024	0.106	0.062	0.857
4	0.393	0.008	0.024	0.135	0.258	6.506	1.995	0.024	0.078	0.036	0.462
5	0.275	0.004	0.024	0.147	0.987	1.705	0.372	0.029	0.075	0.031	0.467
C_1	0.015	-0.003	0.023	0.09	0.189	0.459	-0.052	0.014	0.058	0.031	0.475
C_2	0.007	-0.004	0.032	0.114	0.244	0.586	-0.069	0.013	0.066	0.025	0.308
1x	227.3	6.871	0.021	0.009	0.127	1.342	144.5	0.016	0.012	0.256	0.792
5x	0.137	-0.002	0.023	0.011	0.048	2.124	1.699	0.014	0.005	0.014	0.232
Cx_1	0.013	-0.005	0.021	-0.001	0.018	0.038	0.002	0.011	-0.002	0.008	0.005
Cx_2	0.012	-0.006	0.022	-0.001	0.015	0.041	-0.011	0.01	0.001	0.008	0.016

Table 5

Mean Intensity of Wavelengths Subtracted by Controls (mg/L)											
Samples	As	Cd	Co	Cr	Cu	Fe	Hg	Mo	Ni	Pb	Zn
1	516.489	17.4935	-0.0055	0.024	0.2115	2.0165	49.3205	0.0145	0.011	0.613	1.4375
2	488.789	16.3335	-0.0035	0.057	0.0495	2.1135	41.0405	0.0145	0.035	0.081	0.7755
3	1.291	0.0365	-0.0035	0.095	0.6465	1.0965	7.3825	0.0105	0.044	0.034	0.4655
4	0.382	0.0115	-0.0035	0.033	0.0415	5.9835	2.0555	0.0105	0.016	0.008	0.0705
5	0.264	0.0075	-0.0035	0.045	0.7705	1.1825	0.4325	0.0155	0.013	0.003	0.0755
C_1	0.004	0.0005	-0.0045	-0.012	-0.0275	-0.0635	0.0085	0.0005	-0.004	0.003	0.0835
C_2	-0.004	-0.0005	0.0045	0.012	0.0275	0.0635	-0.0085	-0.0005	0.004	-0.003	-0.0835
1x	227.2875	6.8765	-0.0005	0.01	0.1105	1.3025	144.5045	0.0055	0.0125	0.248	0.7815
5x	0.1245	0.0035	0.0015	0.012	0.0315	2.0845	1.7035	0.0035	0.0055	0.006	0.2215
Cx_1	0.0005	0.0005	-0.0005	0	0.0015	-0.0015	0.0065	0.0005	-0.0015	0	-0.0055
Cx_2	-0.0005	-0.0005	0.0005	0	-0.0015	0.0015	-0.0065	-0.0005	0.0015	0	0.0055

Table 6

Metal Content in Five Herbal Pills in Accordance with Weight and Volume												
Sample	Weight (g)	Volume (ml)	Amount of Metals (mg/kg)									
			As	Cd	Co [2]	Cr	Cu	Fe	Hg	Mo	Ni	Pb	Zn
1	0.5	30	30989.3	1049.6	-0.3	1.4	12.7	121.0	2959.2	0.9	0.7	36.8	86.3
1x [1]	0.36160	50	31428.0	950.8	-0.1	1.4	15.3	180.1	19981.3	0.8	1.7	34.3	108.1
2	0.5	30	29327.3	980.0	-0.2	3.4	3.0	126.8	2462.4	0.9	2.1	4.9	46.5
3	0.5	30	77.5	2.2	-0.2	5.7	38.8	65.8	443.0	0.6	2.6	2.0	27.9
4	0.5	30	22.9	0.7	-0.2	2.0	2.5	359.0	123.3	0.6	1.0	0.5	4.2
5	0.5	30	15.8	0.5	-0.2	2.7	46.2	71.0	26.0	0.9	0.8	0.2	4.5
5x [1]	0.47898	50	13.0	0.4	0.2	1.3	3.3	217.6	177.8	0.4	0.6	0.6	23.1

[1] Samples 1x and 5x are digested by microwave with HNO_3

[2] All cobalt below limit of detection

By observation as samples 1-5 are digested and heated, there is brown vapour in all glass beakers, and the refluxing of brown liquids on the walls of the beakers. Small amount of the vapour leaks out of the gap between the watch glass and beaker. Red solids are present during the heating of the samples in 1 and 2, but are then filtered away.

The wavelength intensity of various metals shown in the controls (C_1, C_2, Cx_1, Cx_2) reflects background noise interference and the possible contamination of metals in the apparatus. Therefore, the mean of the wavelength intensities of C_1 and C_2 has to be subtracted from that of samples 1-5, and Cx_1 and Cx_2 from samples 1x and 5x.

The intensity of cobalt in all samples is below limit of detection. The amount of arsenic, cadmium and mercury are significantly high compared to other metals, particularly in samples 1, 1x and 2. Considerable amounts of copper, iron, and zinc are found in all samples. Less obvious amounts of chromium, molybdenum and nickel are also present. Lead is present in all samples, but relatively more abundant in sample 1.

Sample calculations

1. Intensity of wavelengths subtracted the mean background noise (results in table 5)

$$= Value\ of\ intensity - \frac{(\text{Intensity of control 1} + \text{Intensity of control 2})}{2}\ mg/L$$

Example: intensity of As in sample 1

$$= 516.5 - \frac{(0.015+0.007)}{2}\ mg/L$$

$$= 516.489\ mg/L$$

2. Mineral concentration in mg/kg (results in table 6)

$$= Value\ of\ intensity \times \frac{volume\ of\ sample\ (ml)}{mass\ of\ sample\ (g)}$$

Example: intensity of As in sample 1

$$= 516.489\ mg/L \times \frac{30\ ml}{0.5\ g}$$

$$= 30989.3\ mg/kg$$

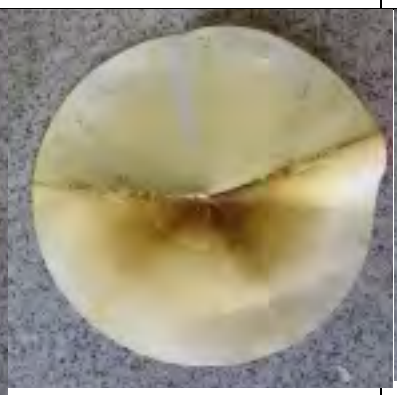
Precipitation as Mercury Sulphide

Table 7 shows the amount of precipitation obtained from filtering the samples. There was little precipitation obtained in samples 'A', thus could not be separated from the filter paper for measurement. It is therefore just compared with filter papers 'B' by observation. All precipitation obtained were brown or yellowish. It is clear that samples 1A, 2A and 3A have more residue on the filter paper, with 1A having particularly dark and thick residue.

It is observed during the boiling of the samples that red powder and indigestible solids are present at the bottom of the beakers containing samples 1 and 2.

Table 7

Precipitated Mercury Sulphide on Filter Papers

Samples	1	2	3
A (with H_2S)			
B (without H_2S)			

Discussion and Evaluation

ICP-OES

Metal content in samples 1-5 by ICP-OES

Metals of Significant Safety Concern - Arsenic, cadmium, mercury and lead

Table 8 and 9 show the limits for toxic elements and heavy metals by the Joint Food and Agricultural Organization (FAO)/ World Health Organization (WHO) Expert Committee on Food Additives (JECFA) and Chinese Medicine Council of Hong Kong (CMCHK) respectively.

The WHO established a *provisional tolerable weekly intake* (PTWI) in 1993, which represents the safe intake contaminants in foods, to prevent the long-term accumulation of contaminants in the human body. The WHO currently does not have a recommended tolerable intake level of arsenic. In 2010, the JECFA mentioned that "the lower limit on the benchmark dose for a 0.5% increased incidence of lung cancer [increase to] 3.0 μg/kg body weight per day" and "noted that the previously established PTWI of 15 μg/kg body weight for inorganic arsenic [in 1989]… was no longer appropriate"(WHO, 2010). Similarly, the previously established PTWI for lead is withdrawn, as it is "not possible to establish a new PTWI that would be considered health protective."

Table 8

Tolerable intake of heavy metals and toxic element JECFA	
Heavy Metal or Toxic Element	PTWI (mg/kg bw) (year established)
As	0.015 (1989) - withdrawn
Cd	0.025 (2013)
Hg	0.004 (2011)
Pb	0.025 (1999) - withdrawn

Table 9

CMCHK Permitted level for heavy metals and toxic element in patent Chinese medicines	
Heavy Metal or Toxic Element	Maximum permitted level
As	1.5 mg/day
Cd	3.5 mg/dosage
Hg	0.036 mg/day
Pb	0.179 mg/day
	(CMCHK 2004)

The content of arsenic in samples is high in all five samples, with samples 1 and 2 being exceptionally high, exceeding PTWI levels by at least 278 mg and 263 mg respectively; CMCHK maximum permitted levels by 268 mg and 253 mg respectively. For example, if a 50 kg person takes pill 1 for three days in a week, one pill (3 g) each time, the amount of arsenic ingested

would be *arsenic content × mass of pill × days = 30989.0/1000 × 3 × 3 =* ***279 mg***, while the tolerable intake per week recommended by JECFA is *body weight × PTWI = 50 × 0.015 =* ***0.75 mg***, meaning a 278 mg exceedance; or if compared to the CMCHK permitted level *permitted level × days = 1.5 × 7* ***= 10.5 mg***, a 269 mg exceedance. These values for other heavy metals can be found in table 10 and in the appendix. Table 10 is produced to compare the amount of heavy metals present to tolerable intake levels of these metals, by converting to amount in mg when a 50 kg individual takes the pill for three days – because some medicines cannot be taken continuously for more than 3 consecutive days – in a week. Please refer to the appendix for the conversion of PTWI and CMCHK limit levels in mg.

Ingesting high amounts of As can lead to arsenic poisoning, interfering mitochondrial respiration, ATP synthesis and ultimately leading to organ failure (Marcus. 2014). Arsenic in these medicines is likely to be from realgar, an arsenic sulphide mineral which has a formula of AsS or As_4S_4. All samples exceed the cadmium limits set by both organisations; again, samples 1 (1049.6 mg/kg) and 2 (980.01 mg/kg) are very high. Although cadmium is not very effectively absorbed in the human body, cadmium can bioaccumulate in organs, the kidneys in particular, leading to renal failure, osteomalacia, osteoporosis and reproduction deficiencies. High amounts of cadmium are also carcinogenic (Anon. 2011). Cadmium is more commonly found in industrial areas and is extremely toxic to humans. Mercury is present in all samples, and is more concentrated in samples 1 and 2. Mercury content in all pills exceeds permitted levels listed above. Cinnabaris (HgS), a component of samples 1 and 2, could be one of the sources of mercury. Mercury is particularly harmful to foetuses as well as infants, putting them at risks of birth defects and neurological dysfunctions due to the inhibited development of the myelin, which might explain why it is recommended not to take these pills during pregnancy, according to the medicine labels. Since mercury attacks the central nervous system, damaging neurons and brain tissues, prolonged mercury exposure could also lead to brain cells dysfunctions along with cognitive, sensory, motor and emotional disturbances (Ralston et al. 2010). Renal failure and acrodynia is also common among those who are frequently exposed to mercury (ATSDR, 1999). Lead is also present in all samples; sample 1 has relatively high lead content of 36.8 mg/kg, but none of the samples exceed tolerable intake levels. Lead occurs naturally in the environment but is also contaminants from industry. It accumulates in the body and results in irreversible neurological, skeletal, renal, cardiovascular harms such as hindering brain development or causing renal impairment (WHO, 2013).

Table 10

The amount of metals ingested if a 50 kg person takes 3 pills a week (mg)											
Sample	As	Cd	Co	Cr	Cu	Fe	Hg	Mo	Ni	Pb	Zn
1	278.904	9.446	-0.003	0.013	0.114	1.089	26.633	0.008	0.006	0.331	0.777
1x	282.852	8.557	-0.001	0.013	0.138	1.621	179.832	0.007	0.015	0.309	0.973
2	263.946	8.820	-0.002	0.031	0.027	1.141	22.162	0.008	0.019	0.044	0.419
3	0.698	0.020	-0.002	0.051	0.349	0.592	3.987	0.005	0.023	0.018	0.251
4	0.206	0.006	-0.002	0.018	0.023	3.231	1.110	0.005	0.009	0.005	0.038
5	0.142	0.005	-0.002	0.024	0.416	0.639	0.234	0.008	0.007	0.002	0.041
5x	0.117	0.004	0.002	0.012	0.030	1.958	1.600	0.004	0.005	0.005	0.208

Key	Exceeds JECFA limits	Exceeds CMCHK limits

Literature value is only available for samples 1, 4 and 5 (Espinoza, et al. 1996). This is shown in table 11. Apart from the amount of arsenic in sample 1, all experimental results are lower than literature values. However, this may not be comparable because of the potential difference in the mixture as a result from time of production and different batches, and that the method in measuring heavy metal content is different – the method used in literature is atomic absorption spectroscopy is instead of atomic emission spectroscopy used in the experiment.

Table 11

Literature analysis of herbal pills using atomic absorption spectroscopy		
Sample	Arsenic range (mg/kg)	Mercury range (mg/kg)
1	1070-12200	26900-207100
4	433-900	7433-60600
5	4967-7367	13867-33000
		(Espinoza, et al. 1996)

Metals of Less Significant Safety Concern – Cobalt, Chromium, Copper, Iron, Molybdenum, Nickle and Zinc

Apart from cobalt, these metals of less significant safety concerns are present in all samples but does not exceed the oral permitted daily exposure (PDE) set by EMEA, as seen in table 12. These metals, if exposed to in small amounts, are harmless. However, above safety levels of exposure to these metals might lead to health problems, which are examined below. The PDE "represents a substance-specific dose that is

Table 12

Metal	Oral Permitted Daily Exposure (mg/kg/day) set by EMEA
Co	N.A.
Cr	0.005
Cu	0.05
Fe	0.26
Mo	0.006
Ni	0.006
Zn	0.026
	(EMEA, 2007)

unlikely to cause an adverse effect if an individual is exposed at or below this dose every day for a lifetime" (EMA, 2012). Please refer to the appendix about conversion of units for comparison with table 10.

Cobalt in the samples is below detection limit. Chromium concentration ranges from 1.4 to 3.4 mg/kg. Chromium (III) exists in foods and water in small amounts and is harmless. Chromium (VI) oppositely is strongly oxidative and changes genetic materials so that they are more likely to mutate, such as defected DNA replication or pairing. Its toxicity could also lead to glomerular and tubular renal injury, hepatic dysfunction and cardiovascular collapse (ATSDR, 2008). Copper content is 2.5-46.2 mg/kg, and are more concentrated in samples 3 and 5. Copper (I) and (II) are essential micronutrients needed for growth and development of connective tissues, bones and various organs, as well as forming red blood cells. On the other hand, free copper is toxic because it interferes with enzymatic reactions and causes oxidative damages to protein, membranes, tissues and organs, especially the liver, kidney and the gastrointestinal tract (ATSDR, 2004). Iron content, 65-359 mg/kg, is comparatively higher in sample 4. Some proteins in the human body contain iron to involve in biological oxidations and transport, such as haemoglobin in the blood, and catalase. It is possible that the high amount of iron in sample 4 is due to its 'vessels reviving properties' i.e. supple ample iron for haemoglobin. Excessive intake of iron however, would produce highly reactive free radicals and causes gastrointestinal damages due to its corrosive nature, resulting in abdominal pain, diarrhea, vomiting, hypovolemia or even organ impairment (Medscape Reference, 2011). Molybdenum is present in small amounts (0.6-0.9 mg/kg). It is considered of low human toxicity, but continuous exposure to Mo could lead to high levels of uric acid in the blood, which could possibly cause kidney stone or develop gout (CDC, 2013). Nickel is present in samples at a range of 0.7-2.6 mg/kg. Nickel is considered very toxic and carcinogenic. Acute ingestion of nickel or its compounds can lead to diarrhea, headaches, shortness of breath, vomiting and cough; while if in large amounts, could cause death (PHE, 2009). And lastly, zinc content in herbal pills ranges from 4.2 to 86.3 mg/kg, with herbal pills 1, 2 and 3 the highest. Zinc catalyses enzymatic reactions and assists the functions of the central nervous system, but when ingested in lethal amounts, zinc chloride is formed when react with hydrochloric acid in the stomach, which is highly corrosive and harms the stomach. Other gastrointestinal distresses include cramps, bleeding, diarrhea and vomiting (ATSDR, 2005).

In view that although the content of these metals is low under the conditions (a 50kg person taking the pill 3 days a week) when compared to tolerable intake levels, the most metal content values are just below the PED values. This means consuming these pills at a higher dosage and regularity for longer periods, is potentially harmful to the individual's health.

To sum up, apart from cobalt, all heavy metals analysed by ICP-OES are present in the herbal pills, with particularly high levels of arsenic, cadmium and mercury. Samples 1 and 2 are the same drug produced by the same manufacturer but in different places. Both pills exceed both tolerable intake levels of JECFA and CMCHK for arsenic, cadmium and mercury. The discrepancies of heavy metal content between the two pills are not large. Herbal pills 1 and 2, produced by Beijing Tong Ren Tang, have a much higher intensity of metals of severe toxicity, than pill 3, a similar drug is produced by Hong Kong Ma Pak Leung. Pill 4 has particularly higher levels of iron, while pill 5 has slightly lower levels of each metal component. All five pills have mercury content exceeding tolerable levels.

Sources of toxic metals

Most herbal pills in China are produce manually and production processes are often badly regulated. Improper production procedures could lead to the contamination of toxic metals in herbal pill components. Besides, due to the unsuitable disposal and treatment of industrial wastes and intensive metal extraction in China, air, soil and water pollution is severe (Dong, et al. 2001). This affects the safety of herbal pills' ingredients, such as herbs grown on contaminated soil, or animal internal organs that accumulate heavy metals ingested from the environment. On the other hand, the presence of heavy metals in the pills could be due to the lack of knowledge of the toxicity of some herbs and minerals, or unauthenticated folk knowledge of components' properties. For example, cinnabaris and realgar are known sources of arsenic and mercury, but are added into the pills because they are believed to have 'reviving properties'.

Comparing wet digestion with HNO_3 to microwave digestion

It is assumed that microwave digestion is more accurate compared to wet digestion as a sealed vessel reduces the leakage of the metals, mercury in particular. Mercury has a lower boiling point of 356.7°C and is more likely to evaporate in the process of heating. This can be seen by the large different Hg intensities in 1, 1x and 5, 5x. However, other metals did not show large intensities differences because they have higher boiling points and do not evaporate as easily.

Errors and improvements in ICP-OES

The possible errors in ICP-OES are contamination of the samples and varying metals intensities of herbal pills.

Only glassware is soaked in nitric acid overnight prior to the experiment, while the funnel and bungs are not. Metals could be present on the funnel and bungs, possibly washed down and mixed into the wet digested samples. To reduce contamination, ensure that all apparatus are soaked before use. Secondly, considering that the herbal pills are manually produced, all components might not be thoroughly mixed and metals intensities could vary within different parts of the herbal pill. Pills from different batches could also vary in metal content. To improve the reliability and representativeness of the results, more samples from different batches should be used, and the mean of metal intensities of each part of the whole pill should be taken. Furthermore, as microwave digestion is more thorough and reduces contamination and leaks, it should be applied to all samples.

An inductively coupled plasma mass spectrometer (ICP-MS) can be used along with ICP-OES, to give a more reliable result.

Precipitation as mercury sulphide

Mercury precipitation is a crude method to identify mercury in the medicine samples, which is not as accurate and significant as ICP-OES. However, it could show a physical amount of mercury present.

Although the three samples are similar in function and content, the amount of mercury precipitation from samples 1A, 2A and 3A varies – 1A is the darkest and densest, while 2A is equally dark but less dense, and 3A is the lightest and has least mercury precipitation. This suggests the varying quality and content of pills produced by each manufacturer in different places. Mercury sulphide is usually red in colour but precipitation on the filter papers tend to be brownish, which suggest the possibility of contamination. This is supported by literature which mentions that "the precipitation of mercury as mercuric sulphide in hydrochloric acid solution is an accurate procedure in the absence of copper, cadmium, tin, zinc and thallium; the latter metals complicate reactions which are based upon the behaviour of pure mercuric sulphide" (Vogel, A. I. 1961). ICP-OES results indicate the presence of copper, cadmium and zinc, which means the formation of mercuric sulphide could be altered by these metals. It was unable to measure the amount of HgS formed as it is difficult to lift HgS off the filter paper and little mercury is formed. This could be due to the incomplete reaction of mercury with hydrogen sulphide or mercury evaporation during the heating process due to its low boiling point. HgS could also be lost into filtrate when the filter paper is not of high enough grade, or that some were filtered out as 'insoluble' parts of the sample. In addition, the indigestible solids present at the bottom of the beakers when boiling samples 1 and 2 could be cinnabaris, and this suggests incomplete digestion of the samples before reaction with H_2S.

To improve this method, the samples can be dissolved in HCl and heated using pressurised vessels and microwave, to ensure that the samples are completely dissolved and mercury does not escape. A higher grade of filter paper can also be used to reduce leakage of HgS into the filtrate.

Conclusion

From both ICP-OES and precipitation of mercury, it is seen that mercury in all pills exceed tolerable intake levels set by both CMCHK and JECFA. Meanwhile, of the other heavy metals examined by ICP-OES, arsenic and cadmium in pill 1 and 2 both exceed JECFA and CMCHK limits. Lead and metals of less significant safety concern does not exceed any tolerable intake limits used in this investigation. From the toxicity of metals examined in the above section, this investigation reflects the need of stricter implementation of quality control of herbal medicines policies and regulations in the usage of these pills, owing that the high amount of mercury in all pills; arsenic and cadmium in pills 1 and 2 in these medicines poses potential threats to users' health.

On the other hand, it is uncertain if the intensity of the range of metals in the pills would 'counteract' each other or result in more complex reactions. Although Chinese medicines are considered pseudoscience, it is still a relatively less explored area of the sciences. Given that it has been quite recognized and has been continuously developing in Asia for over 2000 years, Chinese medicine may contain some truths which are unknown or appear absurd from the western scientific perspective.

Acknowledgements

I would like to thank David Wray and Samantha Booth from the Faculty of Engineering and Science of The University of Greenwich, for arranging and allowing the use of ICP-OES in Chatham campus, in order to complete this investigation.

Bibliography

1. Agency for Toxic Substances and Disease Registry (ATSDR). (1999). *Toxicological Profile for Mercury*. Available: http://www.atsdr.cdc.gov/toxprofiles/tp46.pdf.
2. Agency for Toxic Substances and Disease Registry (ATSDR). (2004). *Toxicological Profile for Copper*. Available: http://www.atsdr.cdc.gov/ToxProfiles/tp132.pdf.
3. Agency for Toxic Substances and Disease Registry (ATSDR). (2005). *Toxicological Profile for Zinc*. Available: http://www.atsdr.cdc.gov/toxprofiles/tp60.pdf.
4. Agency for Toxic Substances and Disease Registry (ATSDR). (2008). *Chromium Toxicity What Are the Physiologic Effects of Chromium Exposure?* Available: http://www.atsdr.cdc.gov/csem/csem.asp?csem=10&po=10. Last accessed 15th July 2014.
5. Anon. (2007). *Ma Pa Leung On Kung Pill.* Available: http://www.mapakleung.com/products_details.php?pid=55&cid=2. Last accessed 18th July, 2014.
6. Anon. (2009). *Tong Ren Tang - Angongniuhuang Pills.* Available: http://www.tongrentang.com/en/medicines/productdetail_1.php. Last accessed 18th July, 2014.
7. Anon. (2009). *Tong Ren Tang - Dahuoluo Pills.* Available: http://www.tongrentang.com/en/medicines/productdetail_2.php. Last accessed 18th July, 2014.
8. Anon. (2009). *Tong Ren Tang –Niuhuang Qingxin Pills.* Available: http://www.tongrentang.com/en/medicines/productdetail_3.php. Last accessed 18th July, 2014.
9. Anon. (2011). *Occupational Safety and Health Administration - Toxic Metals.* Available: https://www.osha.gov/SLTC/metalsheavy/index.html. Last accessed 24th July 2014.
10. Anon. (2013). *Cinnibar in Tongrentang medicines.* Available: http://newsabeta.blogspot.com/2013/06/blog-post_3771.html. Last accessed 11th July, 2014.
11. Centers for Disease Control and Prevention (CDC). (2013). *Biomonitoring Summary – Molybdenum*. Available: http://www.cdc.gov/biomonitoring/Molybdenum_BiomonitoringSummary.html. Last accessed 24th July 2014

12. Chinese Medicine Council of Hong Kong (CMCHK). (2004). *Technical guidelines for registration of proprietary Chinese Medicines – Product Safety Documents Appendix 2.* Available: http://www.cmchk.org.hk/pcm/pdf/guide_save_n_e.pdf.

13. Dong, W.Q.Y., Cui, Y., Liu, X., (2001). *Instances of Soil and Crop Heavy Metal Contamination in China.* Available: http://rmag.ss.msu.ru/articles/448.pdf

14. Espinoza, E.O., Mann, M.-J., Bleasdell, B., DeKorte, S., & Cox, M. (1996). Toxic Metals in Selected Traditional Chinese Medicinals. *Journal of Forensic Sciences, JFSCA*. 41 (3), p453-456.

15. European Medicines Agency (EMA). (2012). *Guidelines on setting health based exposure limits for use in risk identification in the manufacture of different medicinal products in shared facilities.* Available: http://ec.europa.eu/health/files/gmp/2013_gmp_pc_en.pdf.

16. European Medicines Agency (EMEA). (2007). *Guideline on the Specification Limits for Residues of Metal Catalyst.* Available: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003586.pdf.

17. Marcus, S. (2014). *Arsenic Toxicity.* Available: http://emedicine.medscape.com/article/812953-overview. Last accessed 23th Aug, 2014.

18. Medscape Reference. (2011). *Iron Toxicity*. Available: http://emedicine.medscape.com/article/815213-overview. Last accessed 15th July, 2014

19. O'Haver, T. (2014). *Worksheets for Analytical Calibration Curves.* Available: http://terpconnect.umd.edu/~toh/models/CalibrationCurve.html. Last accessed 23th Aug, 2014.

20. Public Health England (PHE). (2009). *Nickel Toxicological Overview.* Available: http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1236757324101.

21. Ralston, N.V.C., Raymond, L.J. (2010). Dietary selenium's protective effects against methylmercury toxicity. *Toxicology*. 278 (1), p112-123.

22. Royal Philips. (2013). *Technical Note 12 - Inductively Coupled Plasma-Atomic Emission Spectrometry (ICP-AES)*. Available: http://www.innovationservices.philips.com/sites/default/files/materials-analysis-icp-aes.pdf.

23. Vogel, A. I. (1961). *Textbook of Quantitative Inorganic Analysis (including elementary instrumental analysis)*. 3rd ed. Great Britain: Longman. p486-487.

24. Wang, X., Zhuang, Z., Sun, D., Hong, J., Wu, X., Lee, F.S.C., Yang, M.S., Leung, H.W. (1999). *Trace Metals in Traditional Chinese Medicine: A Preliminary Study Using ICP-MS for Metal Determination and As Speciation.* Available: http://www.perkinelmer.com/cmsresources/images/44-74381atl_tracemetalsintradlchinesemeds.pdf

25. World Health Organization (WHO). (2005). *WHO Drug Information* 19 (3) Available: http://apps.who.int/medicinedocs/en/d/Js7918e/.

26. World Health Organization (WHO). (2010). *Exposure to arsenic: a major public concern.* Available: http://www.who.int/ipcs/features/arsenic.pdf.

27. World Health Organization (WHO). (2013). *Lead poisoning and health.* Factsheet 379. Available: http://www.who.int/mediacentre/factsheets/fs379/en/. Last accessed 23th July 2014.

28. Zand, N. (2011). *Studies of the nutritional quality of commercial 'ready to eat' infant foods in the United Kingdom*, p103-124 Available: http://gala.gre.ac.uk/8052/1/Nazanin_Zand_Fard_2011.pdf

Appendix

Calculations for comparing the amount of heavy metals ingested

Conditions: a 50 kg person, taking 3 pills a week

A table to show the tolerable intake of heavy metals of significant safety concern for a 50 kg individual taking 3 pills a week

Metals	JECFA – PTWI (mg)	CMCHK – maximum permitted level (mg)
As	0.75	10.5
Cd	1.25	10.5
Hg	0.2	0.252
Pb	1.25	1.253

A table to show the tolerable intake of heavy metals of less significant safety concern for a 50 kg individual taking 3 pills a week

Metal	EMEA – PDE (mg)
Co	N.A.
Cr	0.25
Cu	2.5
Fe	13
Mo	0.3
Ni	0.3
Zn	1.3

Conversion of PTWI – from *mg/kg bw* to *mg*

PTWI (mg/kg bw) × body weight = PTWI (mg)

e.g. PTWI arsenic
= 0.015 mg/kg bw × 50kg
= 0.75 mg

Conversion of CMCHK maximum permitted level – *mg/day* to *mg*, *mg/dosage* to *mg*

Maximum permitted level (mg/day) × days = maximum permitted level (mg)

e.g. maximum permitted level of arsenic
= 1.5 mg/day × 7 days
= 10.5 mg

Maximum permitted level (mg/dosage) × dosage = maximum permitted level (mg)

e.g. maximum permitted level of cadmium
= 3.5 mg/dosage × 3 dosages
= 10.5 mg

Conversion of PDE – mg/kg to mg

PDE (mg/kg) × weight = PDE (mg)

e.g. PDE chromium
= 0.005 (mg/kg) × 50 kg
= 0.25 mg

¿Cómo y en qué formas Lorca presenta el papel de la mujer en *La casa de Bernarda Alba, El amor de Don Perlimplín con Belisa en su jardín y La zapatera prodigiosa?*

Abigail Witts - Spanish

After initially being undecided on whether to do my extended essay in English or Spanish, an essay looking at Spanish literature seemed to be the perfect compromise - and having already studied 'La casa de Bernarda Alba' in class, it did not take me long to decide that I would like to look at more of Lorca's works. The time consuming nature of the process of reading the three books I chose to compare as well as other research papers and literary critiques in Spanish was something I perhaps underestimated, but finding strong links between the three plays I chose was rewarding, enabling me to create a strong argument for my research. The history of Lorca himself was also something that captured my interest, and so I enjoyed being able to look at how his plays were contextualised within this - particularly when it came to how he presented women and society. Whilst at times the task felt overwhelming, I felt a great sense of accomplishment on finishing my essay and found that the more I explored Lorca and his plays, the more I enjoyed them.

Supervisor: Dominic Mott

If the thought of writing a 4000 word essay in a foreign language sounds daunting, then just pause to think quite how much time and effort is required to read the source texts, research the topic, analyse the literary criticism and then plan and formulate the essay itself - all in the target language. That Abie embraced this challenge with open arms speaks volumes for her independence of thought, her determination and her ability to cope with some highly complex and demanding concepts. Entire university courses have been written on Lorca's depiction of the role of women in Spanish society, so for Abie to have forged her own path and developed her own original theory by studying two of his lesser-known texts alongside one of the great classics is quite some achievement. To have written it in such expressive, clear and stylish Spanish is simply the icing on the cake.

Contents

Abstract

Mi monografía consiste en una investigación sobre *cómo y en qué formas Lorca presenta el papel de la mujer en La casa de Bernarda Alba, El amor de Don Perlimplín con Belisa en su jardín y La zapatera prodigiosa*. He dividido mi investigación en dos secciones principales, en primer lugar, atendiendo al comportamiento y el papel de mujeres en las obras y en segundo, más particularmente los recursos poéticos que Lorca utiliza para presentarlas.

Sería imposible examinar todos los poemas y obras de teatro de Lorca, así que decidí centrar mi investigación en tres obras en las que las mujeres son los personajes predominantes.

Mi conclusión es que Lorca crea una presentación coherente de las mujeres a lo largo de las tres obras que he explorado como personajes complejos. Muestra dos tipos de mujeres: las que se conforman, y las que eligen intentar luchar contra las expectativas de la sociedad y su familia. Sin embargo, rebelarse contra la tradición y las expectativas siempre parece traer consecuencias de diferentes tipos para las mujeres. La mayoría de esas consecuencias son malas, pero a veces las consecuencias, aunque vistas como malas por la sociedad, eran preferibles antes de permanecer oprimidos por las propias mujeres.

Al escribir mi ensayo, vi por primera vez las similitudes entre los textos y lo que estas podrían significar en relación con la presentación sobre Lorca de la mujer, decidiendo lo que era relevante y lo que no para mi investigación. Luego, leí lo que algunos críticos había escrito de Lorca sobre los temas que estaba explorando y los usé para desarrollar mi argumento. Además investigué cómo la propia vida de Lorca encajaba con el mensaje de sus obras de teatro, aunque no tuve la oportunidad de explorar esta totalmente en mi investigación.

Introducción

El predomino de las mujeres en las obras de Lorca sugiere que tenía un gran interés en el papel de estas mujeres y lo que representan. Durante la vida de Lorca, las mujeres españolas estaban sujetas a discriminación por parte de la ley y de toda la sociedad, y su opresión es un tema que Lorca explora en las tres obras. Así como las leyes que las dejaron en una posición subordinada a los hombres; la actitud de la sociedad imponía un código de conducta estricto. Debían ser "una madre diligente", "ángeles de la casa" y "la esposa dulce[1]". Sin embargo, estas tres obras fueron escritas entre 1926 y 1936, durante el periodo de la Segunda República cuando las preocupaciones de las mujeres llegaban a ser visibles por primera vez, coincidiendo con la legalización del divorcio y el aborto dando a las mujeres más control sobre sus propias vidas. La época en la que Lorca escribía estas obras era un período de transición donde las mujeres empezaban a hacer frente a la opresión que sufrían; es éste un proceso que el autor explora.

También el propio Lorca era objeto de opresión debido a su sexualidad y ha sido sugerido que es por esta razón que Lorca se identificaba tan grandemente con las mujeres. Así como el homosexual tiene que aguantar las expectativas de los que le rodean y que no tiene ni la inclinación ni capacidad de cumplir, las mujeres, especialmente en las sociedades tradicionales como las que se presentan en las obras de Lorca, tienen que consentir los juicios fundados por los demás y su desprecio. Su deseo por la igualdad y la libertad de la presión para conformarse vincula su

[1] Comenius Project. (2007). *Spanish Women in the 19th Century.* http://www.estelacantabra.com/comenius/SPANISHCENTURY.pdf. Fecha de primer aceso:12.06.2014

sexualidad con su identificación de las mujeres lo que significa que era, según Rafael Martínez Nadal, "un supremo conocedor de la psicología femenina[2]".

Sin embargo, la opresión de las mujeres no se limita al siglo XIX. Una breve mirada de sitio web de El País da titulares de las noticias como "El arquetipo del machista[3]", o "Mujeres inferiores[4]" mostrándonos que el techo de cristal todavía no se ha destrozado. Esto nos muestra cómo la cuestión de la representación y el lugar de la mujer en la sociedad sigue siendo pertinente, lo que significa que podemos seguir aprendiendo de las obras de Lorca y la visión que presenta.

En este ensayo, se razonará si Lorca presenta una actitud coherente respecto a las mujeres en las tres obras a través de la pregunta de investigación '**Cómo y en qué formas Lorca presenta el papel de la mujer en *La casa de Bernarda Alba, El amor de Don Perlimplín con Belisa en su jardín* y *La zapatera prodigiosa*?**' Las mujeres en estas obras tienen dos opciones ante esta opresión: someterse a este papel sin conocimiento completo de que serán conducidas a su propia tristeza, como Adelaida en "La casa de Bernarda Alba" cuyo prometido no la deja salir de casa, o pueden intentar superar la opresión, aunque hacer esto tiene un precio. Los personajes principales que luchan contra la opresión son Bernarda, a precio de su feminidad, Adela, a costa de su vida, la zapatera, a costa de su lugar en la sociedad y Belisa, a costa del hombre a quien pensaba que amaba, dejándola en una búsqueda interminable de su amante misterioso.

[2] Loureiro, A. 1988. *Estelas, laberintos, nuevas sendas*. Barcelona: Gráficas Alpes, página 264.

[3] Sanchez, V. 2014. *El arquetipo del machista*, http://elpais.com/elpais/2014/05/16/opinion/1400260086_326171.html. Fecha de primer aceso: 01.07.2014.

[4] Torres Lopez, J. (2014). *Mujeres inferiors*, http://ccaa.elpais.com/ccaa/2014/05/18/andalucia/1400418796_715217.html. Fecha de primer aceso 18.08.2014.

En primer lugar, observaré cómo Lorca utiliza lo que las mujeres hacen y las situaciones en las que se encuentran, junto con el contexto histórico, para explorar su papel. A continuación, examinaré más específicamente los recursos poéticos y teatrales utilizados por Lorca para reforzar su descripción de la función de la mujer en la sociedad.

El papel claramente definido de las mujeres

Las obras muestran que el papel de las mujeres en la sociedad estaba claramente definido y subordinado al de los hombres – un papel que, según un estudio sobre las mujeres en España durante el siglo XX, estaba garantizado por "el sistema patriarcal"[5]. Lorca retrata las sociedades en estas tres obras como opresor hacia las mujeres, reforzando este papel de subordinación, y la opinión de Amelia que "Nacer mujer es el mayor castigo[6]" parece establecer el tono de cómo las mujeres en todas las obras son encarceladas por lo que se espera de ellas.

Lorca establece cuál es el lugar de la mujer en el hogar a través de la afirmación de Bernarda que es "Hilo y aguja para las hembras[7]", "Látigo y mula para el varón[8]". Esta yuxtaposición de papeles hace hincapié en que las mujeres debían cumplir y vivir dentro de límites muy estrechos que dictaban lo que debían hacer. Sin embargo, en 'La casa de Bernarda Alba' vemos a Bernarda desafiar este estereotipo, organizando su finca como un hombre y manteniendo el orden con sus gritos y golpes de su

[5] Martinez, A (2008). *La situación de la mujer a principios del siglo XX*. Espana: uji. Pagina 7

[6] García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 48

[7] García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 15

[8] García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 15

bastón. A través de este conflicto entre la descripción de lo que la mujer debe ser y cómo las mujeres como Bernarda en realidad se comportan, Lorca empieza a mostrar la presencia de la resistencia de las mujeres contra su papel.

Las situaciones de la zapatera y Belisa también muestran cómo las mujeres eran vistas como dependientes de los hombres. La madre de Belisa enfatiza que Don Perlimplín "tiene muchas tierras[9]", sugiriendo que Belisa depende de casarse con él por razones prácticas. Además, en realidad, la zapatera no hace zapatos, pero toma la identidad de su marido. Sin embargo, esta dependencia de él se muestra como innecesaria cuando su marido la deja y ella trabaja apoyándose a sí misma, con lo cual Lorca sugiere que estos papeles restrictivos no son necesarios por ninguna razón.

El matrimonio y las relaciones con los hombres

La presión sobre las mujeres para casarse con la persona que los padres habían escogido está presente en las tres obras. Belisa no tiene opción en su matrimonio con Don Perlimplín, sus súplicas a su madre para tener voz y voto se encuentran con la afirmación brusca: "Tu estás conforme, naturalmente[10]"

Este tema de conformidad también está presente en "La casa de Bernarda Alba" y "La Zapatera prodigiosa". Adela no puede casarse con el hombre que ama porque ya se ha decidido que se casará con su hermana, Angustias, y la infelicidad en el matrimonio de la zapatera está en primer plano desde el inicio de la obra cuando dice, entre

[9] García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 257

[10]García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 257

lágrimas, “Quién me hubiera dicho a mi… que me iba a ver casado con…! Me tiraría del pelo.[11]”. Su infelicidad sugiere que no hay una posibilidad de escapar de la relación, creando una sensación de atrapamiento.

En ‘La casa de Bernarda Alba’, Adela habla de “Santa Bárbara bendita”. En el cuento de Santa Bárbara, su padre construyó una torre muy alta y allí “encerró a su hija hasta que decidió entregarla en matrimonio a uno de los príncipes que la pretendían atraídos por el misterio que rodeaba su encierro[12]”. Santa Bárbara escapó de la torre en la ausencia de su padre y por consiguiente la mató a su regreso. En el caso de “La casa de Bernarda Alba”, la referencia a esta historia anticipa directamente la muerte de Adela que viene de su búsqueda por la libertad. Sin embargo, este ‘precio’ es un tema también evidente en las otras dos obras, y la falta de elección recuerda a los matrimonios forzados de la zapatera y Belisa.

La opresión de las mujeres por las mujeres

Con frecuencia, Lorca también muestra cómo las mujeres son cómplices en la opresión de otras mujeres en la sociedad. Como he explorado, en “El amor de Don Perlimplín”, es la madre de Belisa la que es fundamental en la organización del matrimonio forzado de su hija, sin pensar en lo que ella desea. Esta influencia de la sociedad exterior en las familias hace hincapié en la universalidad de la opresión, que

[11] García Lorca, F. 1973. La zapatera prodigiosa. Barcelona: Espasa Libros, página 62

[12] Juanca, A. (2008). *Santa Bárbara Virgen y Mártir.* http://santabarbara.co.cr/notas.item.12/santa-barbara-virgen-martir.html. Fecha de primer aceso: 14.08.2014.

creó, según un socialista prolífico en el momento en que Lorca estaba escribiendo, una "cultura de opresión femenina[13]".

La opresión de las mujeres no se limita a los que están dentro de sus propios hogares. En "La zapatera prodigiosa", en lugar de apoyarla a ella, las vecinas de la zapatera "adoptan una actitud cómica de pena[14]" y la atormentan cuando su esposo la deja, mostrando su crueldad.

Este aislamiento de la mujer en la sociedad también esta presente en "La casa de Bernarda Alba" cuando Martirio habla sobre el deseo de unirse al linchamiento de "la soltera [que] tuvo un hijo no se sabe con quien[15]". Martirio dice "que pague lo que debe[16]", mostrando cómo, en vez de tener simpatía por una mujer de su misma edad y clase, mantiene los valores que su familia y la sociedad le han enseñado.

Las consecuencias de no conformar

En las tres obras, las mujeres son aceptadas si se conforman, pero para aquellas que optan por no conformarse hay consecuencias inevitables – la mayoría de las cuales son negativas.

[13] Porta, M. (1982). *Mujeres e industrias culturales.*. Available: http://www.mujeresenred.net/spip.php?article693. Fecha de primer aceso 17.09.2014.
[14] García Lorca, F. 1973. La zapatera prodigiosa. Barcelona: Espasa Libros, página 97
[15]García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 65
[16]García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 65

Para Belisa y Bernarda, la consecuencia es el rechazo de la opinión tradicional de la feminidad, aunque esta pérdida de la feminidad no es completamente negativa para las mujeres, las identifica como diferentes de los demás en la sociedad .

Según una investigación sobre la mujer española a finales del siglo XIX por Teresa Gómez Trueba, la mujer debía tener "pasividad, sentimiento, fragilidad y su función estaba en el hogar[17]"pero Bernarda desafía este estereotipo con violencia cuando grita y da "*un golpe de bastón en el suelo*[18]", casi asumiendo el papel de un líder religioso con el canto misterioso en el funeral de su segundo marido. Bernarda se presenta como un personaje masculino, enfatizado aún más por la falta de personajes masculinos, mostrándose como no conforme a las expectativas de las mujeres a ser pasivas y, la ha llevado a perder su feminidad por completo. Ser diferente a lo que se espera también provoca que Bernarda sea rechazada por la sociedad; incluso la familia de su marido "la odia[19]", según la Poncia.

Como Bernarda, Belisa también pierde su sentido de la feminidad y la calidad humana como personaje. Empezamos a ver su rebelión en cómo una esposa 'debe' comportarse a través de referencias de Lorca a su tono de voz como "mimosa" y "guasona[20]" cuando habla a Don Perlimplín. Aunque en este momento en la obra ella todavía es exteriormente sumisa a su marido, podemos sentir el descontento hacia él. Aunque su renuncia a las expectativas de la sociedad son menos obvias, su presentación como un personaje puro e inocente en el comienzo de la obra portando

[17] Gomez Trueba, T. (1988). Imagenes de la mujer en la España de finales del siglo XX . En: Garcia Carcel, R *La mujer en Espana* . Espana: Universidad de Valladolid. Pagina 10.

[18] García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 11

[19] García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 5

[20]García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 263

un vestido "lleno de encajes[21]", en contraste con su presentación en el final de la obra donde el encaje puro está cubierto en la sangre de Don Perlimplín, muestra cómo su pureza se ha convertido en corrupta, que la aparta de la imagen tradicional de la feminidad.

Mientras que a muchos de los espectadores de Lorca el momento de pérdida de la feminidad habría sido visto como algo negativo, el hecho de que lo presenta más ambiguamente (con mujeres como Bernarda casi persiguiendo un papel masculino) sugiere que Lorca está en parte tratando de desafiar el concepto de la feminidad definido por la sociedad.

También podemos ver semejanzas entre el precio que tanto la zapatera como Adela pagan por no conformarse. Adela lucha contra la expectativa que he explorado previamente acerca de casarse con quien se le dice, desafiando no sólo el papel tradicional de una mujer, sino también las órdenes de su madre. Sin embargo, la primera muestra que obtenemos de la naturaleza rebelde de Adela no viene de sus acciones, sino del nombre que Lorca le da. El nombre 'Adela' "viene del verbo español 'adelantar'[22]", según un estudio del uso de Lorca de nombres en la obra. Adela está luchando para adelantarse a la opresión frente a ella.

En el acto tercero, después de una discusión acalorada con Martirio, Adela dice "todo el pueblo contra mí[23]", destacando una sensación de aislamiento que se refleja en los

[21] García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 261

[22] Anon. (2012). *La casa de Bernarda Alba, personajes.* Available: http://centrodeartigos.com/articulos-de-todos-los-temas/article_20286.html. Fecha de primer aceso 15.07.2014.

[23]García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 87

últimos momentos de 'La Zapatera prodigiosa' cuando la zapatera declara como "ya somos dos [ella y su marido] a defender mi casa[24]". Ambos discursos muestran cómo se puede rebelarse – o por lo menos intentarlo – pero si hace esto existe el riesgo de aislamiento y de rechazo de la sociedad.

Además, Lorca también sugiere que rebelarse supone un precio más alto. En última instancia, la rebeldía de Adela la lleva a su suicidio cuando ella cree que su Pepe ha sido disparado como consecuencia de los problemas causados por sus relaciones. Lo mismo sucede cuando las acciones de Belisa llevan directamente a la muerte de su marido. Esto muestra a la audiencia que las mujeres que Lorca está retratando no se rebelan sin causa. Las posibles consecuencias significan que la decisión de luchar contra las expectativas no se debe tomar a la ligera.

Análisis de los símbolos

Hasta ahora he explorado lo que podemos deducir de la presentación de Lorca de las mujeres y su comportamiento. Ahora exploraré cómo las técnicas literarias que utiliza dan una visión más sutil de los personajes.

La presencia del verde

Lorca usa color en las tres obras para reforzar su presentación del papel de las mujeres creado por sus acciones, situaciones e interacciones de una manera más sutil y poética. Una de las principales formas en que hace esto es a través de la ropa de las

[24] García Lorca, F. 1973. La zapatera prodigiosa. Barcelona: Espasa Libros, página 155

mujeres – especialmente Adela, la zapatera y Belisa. Aunque hay muchos colores que me hubiera gustado explorar, me voy a concentrar en el color más prominente – el verde.

Barbara Hranáčová, en su análisis de simbolismo en 'La casa de Bernarda Alba', escribió que "Debido a los significados contrarios que le representa Lorca, el verde es uno de los colores más problemáticos de su obra[25]" y este carácter problemático del verde también está presente en las otras dos obras. Hranáčová explica como, por un lado, puede simbolizar la vida y la vitalidad, pero por otro lado se usa muy frecuentemente en la poesía de Lorca para simbolizar la muerte. A la vez que problemática, su prevalencia en todas las obras de Lorca indica que es un color muy simbólico.

El uso de verde parece simbolizar una rebelión contra la sociedad, el color vivo luchando contra la oscuridad de las expectativas y por lo tanto es significativo que tanto la zapatera como Adela lleven vestidos verdes en algún momento en cada obra. El vestido verde que la zapatera lleva al principio, un "verde rabioso[26]", subraya la violencia del color mostrando su pasión y vivacidad y otra manera en que la zapatera está luchando contra la expectativa de ser pasiva. Adela también lleva un vestido verde en un contexto similar: después de la muerte del marido de Bernarda, la familia debe llevar negro, pero en un momento de rebelión, Adela declara "¡Mañana me pondré mi vestido verde y me echaré a pasear por la calle![27]". La reacción de

[25] Hranacova, B. (2009). *El simbolismo en La casa de Bernarda Alba.* Available: https://is.muni.cz/th/201754/ff_b/bc.txt. Fecha de primer aceso 07.07.2014.

[26] García Lorca, F. 1973. La zapatera prodigiosa. Barcelona: Espasa Libros, página 61

[27] García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 29

Martirio, haciendo callar a Adela "con autoridad[28]", demuestra que es una de esas mujeres, como ya hemos visto antes, que contribuye a la opresión de la mujer en la sociedad.

Mientras que en 'La casa de Bernarda Alba' y 'La zapatera prodigiosa' son las mujeres las que llevan el verde, en 'Amor de Don Perlimplín', es el propio Don Perlimplín. La obra comienza con él llevando "una casaca verde[29]", y el color es intensificado por "las paredes verdes[30]" de su casa. En este contexto, no parece que podamos inferir el mismo significado que cuando se utiliza para los vestidos de Adela y la zapatera. En primer lugar, podemos observar la asociación más tradicional del verde con la envidia y los celos. Con estas connotaciones, parece que Lorca está usando el verde para demostrar el efecto que Belisa ha tenido en él – el hecho de que Perlimplín esta corroído de envidia evidencia el éxito de Belisa desafiando su papel prescrito como una esposa fiel y pasiva.

En 'El amor de Don Perlimplín', el verde simboliza la esperanza falsa que tenía por su matrimonio, que es arruinado por la infidelidad de Belisa y los celos que esta causa. Si combinamos estos significados con las connotaciones contrastantes de la vida y de la muerte explicadas por Hranáčová, Lorca parece estar utilizando el verde para indicar el crecimiento o comienzo de una rebelión contra la sociedad y posiblemente el fin del cumplimiento del estereotipo impuesto sobre ellas - y el conflicto que esto causa - y refuerza la idea de las posibles consecuencias negativas que una rebelión puede traer.

[28] García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 29

[29] García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 253

[30] García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 263

El uso de las flores

Lorca utiliza las imágenes de las flores para presentar y explorar el papel de las mujeres en sus obras, y todas estas referencias refuerzan la imagen tradicional de las mujeres, dándoles los atributos de flores delicadas que necesitan ser cuidadas y nutridas.

El alcalde en 'La zapatera prodigiosa' comenta cómo muchas mujeres que ha conocido son "como amapolas[31]" o "como rosas de olor[32]", seguido más tarde en la obra por el zapatero diciendo a su esposa que ella es una "clavelinita encarnada[33]". Estas referencias a las mujeres como flores delicadas también están presentes en 'El amor de Don Perlimplín', cuando Marcolfa dice a Belisa que "ella se puso encendida como un geranio[34]" y cuando, hablando a Don Perlimplín sobre el matrimonio de su hija, la madre de Belisa le asegura que ella "es una azucena[35]" – una flor asociada con la inocencia y la pureza.

Así como demostrando la inocencia esperada de las mujeres, las flores en las tres obras parecen mostrar el crecimiento de la lucha de las mujeres contra las expectativas de la sociedad. En 'La casa de Bernarda Alba', la Poncia cuenta la historia de Paca la Roseta, quien fue secuestrada y llevada por un grupo de hombres

[31] García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 111

[32] García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 111

[33] García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 137

[34] García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 281

[35] García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 257

"tan conforme[36]". En su regreso llevaba "una corona de flores en su cabeza[37]". Este uso de las flores, combinado con su "pelo suelto[38]", parece representar la libertad que ella ha experimentado y, de la misma manera que las flores evocan imágenes de la naturaleza, ella ha hecho lo que quería y ha renunciado a las restricciones que le habían sido impuestas. También la zapatera se describe con "pelo tirante, adornado con dos grandes rosas[39]". De la misma manera que hay un conflicto en el uso del verde en la obra de Lorca, su uso de las flores muestra un conflicto entre la forma inocente en la que se espera que las mujeres se comporten y la forma en que se ven cuando se rebelan.

La presencia de los animales

Lorca utiliza imágenes de animales para demostrar el papel asignado a las mujeres por la sociedad – como artero, loco, con un propósito claro, ya sea como extremadamente pasivo o dominante. En 'La casa de Bernarda Alba', María Josefa canta que Magdalena tiene "la cara de hiena[40]". A esto hace referencia la vecina amarilla en 'La zapatera prodigiosa' cuando describe la casa de la zapatera como "de una hiena[41]". En algunas culturas, la hiena puede simbolizar "la desviación sexual[42]" o se ve como un "animal malvado y sucio[43]" y por eso Lorca parece demostrar como las mujeres

[36] García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 19

[37] García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 19

[38] García Lorca, F. 1990. Amor de Don Perlimplin con Belisa en su jardín, Madrid: Ediciones Cátedra, página 19

[39] García Lorca, F. 1973. La zapatera prodigiosa. Barcelona: Espasa Libros, página 61

[40] García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 82

[41] García Lorca, F. 1973. La zapatera prodigiosa. Barcelona: Espasa Libros, página 143

[42] Anon. (2011). *Signos y simbolos* . Available: http://signosysimbolos.wordpress.com. Fecha de primer aceso 12.07.2014.

[43] Anon. (2014). *Simbologia animal.* http://www.circuloromanico.com/index.php?menu_id=5&jera_id=2355&page_id=1871&cont_id=5059. Fecha de primer aceso 23.07.2014.

etiquetan como astutas y arteras a otras mujeres que no adaptarse a cómo piensan que deben comportarse.

Al lado de las comparaciones con las hienas, Lorca también utiliza un perro en su presentación de las mujeres. Martirio dice que lo que les importa a los hombres es tener "una perra sumisa[44]", una frase que parece resumir cómo los hombres perciben a las mujeres en las obras. Sin embargo, no solo son las mujeres las que Lorca presenta como débiles. En 'La zapatera prodigiosa' es el hombre quien es comparado con un perro cuando el zapatero dice que es "un perillo y mi mujer manda en el castillo[45]", mostrando como, cuando una mujer, como su esposa, juega un papel más dominante, ella todavía no puede escapar de ser etiquetada negativamente – en este caso, el zapatero parece estar insinuando que, por ser más dominante, su mujer le ha emasculado.

Además podemos ver otros casos de mujeres que son etiquetadas como animales, por ejemplo como locas, cuando la madre de Belisa se describe con su peluca "llena de pájaros" o María Josefa al hablar con su oveja. Parece ser que, a través de la comparación de las mujeres con los animales, Lorca está mostrando la forma en que se agrupan y son estereotipadas por la sociedad, y, al traer estos estereotipos a nuestra atención, él está tratando de denunciarlos.

[44] García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press, página 23

[45] García Lorca, F. 1973. La zapatera prodigiosa. Barcelona: Espasa Libros, página 152

Conclusión

En este ensayo he intentado contestar la pregunta '**Cómo y en qué formas Lorca presenta el papel de la mujer en *La casa de Bernarda Alba, El amor de Don Perlimplín con Belisa en su jardín* y *La zapatera prodigiosa*?**', y llego a la conclusión de que Lorca presenta dos tipos de mujeres en estas tres obras. Por una parte, están las que se ajustan a los límites del papel que la sociedad les impone. Las hermanas de Adela, los otros personajes femeninos en 'La zapatera prodigiosa' y la madre de Belisa no muestran ninguna señal de intentar cambiar su lugar en la sociedad, destacando que no todas las mujeres lo hacen. Es difícil saber lo que hace que estas mujeres hagan esto, ya que estas mujeres no se exploran en tanta profundidad por Lorca, pero es evidente que la decisión de ajustarse no garantiza la felicidad, aunque podría decirse que conduce a una vida más aparentemente tranquila y silenciosa. Por otra parte, están las que tratan de luchar contra el papel que la sociedad les confina, pero con consecuencias. Como he explorado, estas 'consecuencias' no son uniformes ni necesariamente predecibles, pero renunciar – de cualquier manera – a las expectativas a las que las someten, es posible, en cierta medida, aunque no se presenta como una opción fácil.

Esto crea un conflicto en la presentación de Lorca de las mujeres: algunos pueden o desean liberarse de las expectativas y límites de ser mujer, mientras que otros permanecen confinadas por la tradición.

También esta claro que el método de Lorca para la creación de esta presentación del papel de la mujer puede serpararse en dos secciones principales: por una parte, las

acciones de las mujeres y su entorno social, y por otra, las técnicas poéticas más sutiles que nos insinúan su carácter. No sabemos mucho sobre las opiniones reales de Lorca, por lo que tenemos que interpretar sus obras con el fin de tratar de entenderlo. Al observar en primer lugar las acciones de las mujeres y las situaciones en que se encuentran, podemos ver los cuatro argumentos principales que Lorca parece estar presentando sobre el papel de la mujer. Primero, que su papel en la sociedad está muy claramente definido y restringido; segundo, que su elección en las relaciones es limitada; tercero, que los hombres no son los únicos culpables de la opresión sexista y, finalmente, que normalmente hay un precio que pagar por no conformarse. Parece que el uso del verde junto al imaginario de las flores y los animales aún refuerza más estos cuatro puntos, dando una profundidad a los personajes. No solo da las obras la tradicional sensación poética de Lorca, sino que también nos permite ver una vez más a las mujeres como gente mucho más compleja. Si hubiera sido posible dentro de mi límite de palabras, me hubiera gustado haber investigado más sobre lo que sabemos acerca de las opiniones reales de las mujeres y la opresión de Lorca y compararlo con lo que podemos deducir de sus obras. A lo largo de su obra, Lorca intenta socavar los estereotipos con los que las mujeres fueron marcadas, sin embargo, el hecho de que aún existe la opresión de las mujeres casi cien años después sugiere que su mensaje sigue siendo hoy tan pertinente como entonces.

Bibliografía

PÁGINAS WEB

Anon. (2011). Signos y simbolos . Available: http://signosysimbolos.wordpress.com. Fecha de primer aceso 12.07.2014.

Anon. (2012). La casa de Bernarda Alba, personajes. Available: http://centrodeartigos.com/articulos-de-todos-los-temas/article_20286.html. Fecha de primer aceso 15.07.2014.

Anon.(2014) Simbologia animal. http://www.circuloromanico.com/index.php?menu_id=5&jera_id=2355&page_id=1871&cont_id=5059. Fecha de primer aceso 23.07.2014.

Aspinall, H. (2012). The Fetishization and Objectification of the Female Body. http://arts.brighton.ac.uk/study/english-literature-studies-brighton/brightonline/issue-number-two/the-fetishization-and-objectification-of-the-female-body-in-victorian-culture. Fecha de primer aceso 05.07.2014.

Comenius Project. (2007). Spanish Women in the 19th Century. http://www.estelacantabra.com/comenius/SPANISHCENTURY.pdf. Fecha de primer aceso:12.06.2014

Porta, M. (1982). Mujeres e industrias culturales.. http://www.mujeresenred.net/spip.php?article693. Fecha de primer aceso 17.09.2014.

Sanchez, V. 2014. El arquetipo del machista, http://elpais.com/elpais/2014/05/16/opinion/1400260086_326171.html. Fecha de primer aceso: 01.07.2014.

Torres Lopez, J. (2014). Mujeres inferiors, http://ccaa.elpais.com/ccaa/2014/05/18/andalucia/1400418796_715217.html. Fecha de primer aceso 18.08.2014.

LIBROS

Bonaddio, F (2010). Federico Garcia Lorca. London: Boydell & Brewer

García Lorca, F. 1990. Amor de Don Perlimplín con Belisa en su jardín, Madrid: Ediciones Cátedra

García Lorca, F. 1983. La casa de Bernarda Alba. Manchester: Manchester University Press

García Lorca, F. 1973. La zapatera prodigiosa. Barcelona: Espasa Libros

Gomez Trueba, T. (1988). Imagenes de la mujer en la España de finales del siglo XX . En: Garcia Carcel, R La mujer en Espana . Espana: Universidad de Valladolid.

Loureiro, A. 1988. Estelas, laberintos, nuevas sendas. Barcelona: Gráficas Alpes,

Martinez, A (2008). La situación de la mujer a principios del siglo XX. Espana: uji.

Celebrating the achievements of Sevenoaks students

Sevenoaks School was one of the first schools in the UK to introduce the International Baccalaureate, and the first UK school to move exclusively to the IB at Sixth Form. The school now enjoys world-class academic status, with an average IB score ten points above the world average.

The IB has rapidly established itself as the international gold standard of education. Its single vision is an attempt to prepare young people for a life of learning in an unpredictable future. While having a proven track record in its ability to prepare students better for undergraduate studies, its educational vision is premised on a notion of what being well-educated means: an 18-year-old school-leaver should be numerate, literate, able to apply rigorous scientific and mathematical thinking to the world, while being able to engage in foreign cultures. These are the basic requirements of living in the multifaceted world of the 21st century, regardless of the university destination or degree.

In this book, we have a collection of ten of our IB Extended Essays from our 2015 cohort of students – these typify the depth, breadth and subject nature of the studies that form an integral part of the IB.

We hope you enjoy reading them!

Sevenoaks School
High Street, Sevenoaks, Kent TN13 1HU

www.sevenoaksschool.org

directdesign.co.uk