



A Psychological Model of Translation Quality Assessment: Applying Bloom's Taxonomy of Educational Objectives to Grade English Texts

Natasha Pourdana (Corresponding author)

Department of English Translation, Faculty of Literature and Foreign Languages,
Islamic Azad University, Karaj Branch, Iran

&

Global Education Center, Gyeongju University, South Korea

E-mail: Natasha.qala@gmail.com

John S. Rajeski

E-mail: asiajsr@gmail.com

Received: 08-11-2013

Accepted: 09-12-2013

Published: 01-03-2014

doi:10.7575/aiac.ijalel.v.3n.2p.89

URL: <http://dx.doi.org/10.7575/aiac.ijalel.v.3n.2p.89>

Abstract

The present study explored the possibility of grading English texts based on the psychological processes a non-native English translator might go through by applying Benjamin Bloom's Taxonomy of Educational Objectives: Cognitive Domain. The researchers' primary assumption was that the difficulty of English texts would increase as the translator's performance proceeds from translating an English text that requires their only language Knowledge through the texts requiring increasing demands of their Comprehension, Application, Analysis, Aynthesis and Evaluation abilities. To the researchers, the quality of the translation performance could be assessed based on the type of psychological processing they conduct in translating a given text. To fulfil the ultimate purpose of the research, 30 undergraduate students majoring in English translation at Islamic Azad University, Karaj Branch, Iran took part in this investigation in 2012. In addition to Comprehensive English Language Test (CELT), the participants were given a translation essay-type test that included six short English passages of 150 to 200 words. The texts were carefully graded based on the type of behavioural objectives (i.e., psychological processes) demanded of the translator and listed according to the hierarchy of Bloom's cognitive domain. A non-parametric test of Kruskal Wallis proved that except for the fifth level, the Synthesis text, the trainee translators' quality of performance was graded properly based on the difficulty levels expected and explored in Bloom's levels of cognitive domain. The findings in this study are considerably practical in translation quality assessment, upgrading the reliability estimate of translation tests and teaching translation skills and strategies.

Keywords: EFL, Bloom's Cognitive Domain, Grading, Psychological Processing, Translation Quality Assessment

1. Introduction

Throughout history, written and spoken translators have played a crucial role in intercultural or cross-cultural (the latter term would be my initial choice) communication, not least in providing access to important texts for scholarship and/or religious purposes. Yet the study of translation as an academic subject only really began during the past sixty years. In the English-speaking world, this discipline is now generally known as *Translation Studies*. This nascent discipline is mostly concerned with 'the complex of problems clustered round the phenomenon of translating and translation' (Holmes, 1988b:181). Translation studies are known as an academic research area that has expanded exponentially in recent years.

Translation is formally studied as a language-learning methodology or as a part of comparative literature, translation workshops and contrastive linguistics courses. However, the discipline as it is now known owes much to the seminal work of James S. Holmes, whose 'The name and nature of translation studies' (1988b/2004) proposed both a name and a structure to the field (Munday, 2001). Crucially, Holmes puts forward an overall framework, describing what translation studies entail. In Holmes' explanations of this framework, a branch of 'pure' research in translation is descriptive (2004:184-90). Descriptive translation studies have three possible foci: examination of (1) the product, (2) the function, and (3) the process of translation.

- (1) Product-oriented study examines existing translations. One of the objectives of the product-oriented studies might possibly be general history of translations – however ambitious such a goal might sound at this time.
- (2) Function-oriented study describes the 'function of translation in the recipient sociocultural situation: it is a study of contexts rather than texts'.
- (3) Process-oriented study is concerned with 'the psychology of translation or with trying to find out what happens in the mind of a translator'. Despite later work from a cognitive perspective including think-aloud protocols

(where recordings are made of translators' verbalization of the translation process as they translate), this is an area of research which is only now being systematically analyzed (Munday, 2001: 10-11).

The psychological modelling of translation process, as Campbell (1999) states, is done to infer mental constructs from empirical data. The main examples are the works of think-aloud methodologists such as those reported in Faerch and Kasper (1987). In psychological modelling, the attempt is made to map the mental constructs in translators at different levels of ability, or longitudinally in student translators (Campbell, 1999).

Two comprehensive psychological modeling(s) are proposed in Mason (2001) and Krings (1986). Mason sets out to investigate translation strategies in oral translation into a second language. From think-aloud data a range of strategies are detected, including 'realizing a translational problem', 'monitoring Source Language- or Target Language- text segments', and 'rephrasing Target Language- text segments' (Munday, 2001:7). Similarly, Krings develops a flow chart to illustrate his 'tentative model of the translation process' (1986:269). This includes such strategies as retrieval, monitoring and decision-making.

On the other hand, Roger Bell (1991:43) in his Interpretive model of translation asserts that 'focus on the description of the process and/or the translator form the twin issues that translation theory must address: how the process takes place and what knowledge and skills the translator must possess in order to carry it out'. Thus, the Interpretive model of translation explains translation as an overlapping three-stage process involving:

- (1) Reading and understanding: Using linguistic competence and 'world knowledge' to grasp the sense of the source text. The linguistic component needs to be understood by reference not only to explicit but also to implicit meaning in an attempt to recover the authorial intention. The world knowledge we have is de-verbalized, theoretical, general, encyclopedic, cultural and activated differently by different translators and by different texts:

'Translators are privileged readers called on to understand the facts in a text and to feel its emotional connotations. That is why translators do not feel equally close to all texts' (Lefevere, 1993:31).

- (2) De-verbalization: 'an essential intermediate phase if the translator is to avoid transcoding and calques' (p.115). An explanation developed to explain the cognitive processing of the interpreter, where transfer is supposedly through sense and not words, de-verbalization is claimed to be less obvious in translation but still is present.
- (3) Re-expression: (Lefevere, 1993:132) where the target text is constituted and given form based on the de-verbalized understanding of sense.
- (4) Verification: where the translator revisits and evaluates the target text.

What they all have in common is the empirical data, which are used to infer mental constructs about translation; there is no reason why a model of assessing translation competence cannot be built around these kinds of construct.

Translator and interpreter education is now widely practiced around the world and is supported by an increasingly sophisticated body of research and scholarship. Much of this work is concerned with identifying the components of competence and proposing curriculum models that incorporate these components and suitable teaching strategies. The scholarship supporting translation and interpreting education necessarily entails discussions of assessment and there has been some encouraging work in this area.

The translation of a text appears to be the standard form for translation quality assessment in educational settings, although Pym (2004b) describes such tasks as skimming and scanning materials in English and producing 'written commentaries in the other language' (2004b:53). Little is written about the basic approaches of translation instruments, therefore, assessing the relative achievement of the top scoring trainee translators; depending on where the bottom of scale finds itself, one may be very good and one exceptional, or perhaps they are separated by a whisker margin. Moreover, very little is written on interpreting assessment. This may be partly due to the relatively few formal courses in the field worldwide, to the limited research in the area, and to the intuitive nature of the oral translation test design and assessment criteria.

Common to all discussions in the literature are the competencies that translation tests aim to assess, the subjective criteria, and the high failure rate (Munday, 2008). There is some general agreement on the skills and abilities necessary for a student translator to succeed on an oral translation course or in the profession (Lambert, 1991), although this is not based on any empirical data, but rather on intuitive judgments by translator trainers. These competencies include: 'good command of the relevant languages, speed of comprehension and production, good general knowledge of the world, good memory, test-taking strategies, and stress tolerance' (1991:39). Translation tests tend to be criterion-referenced, with candidates required to reach each criterion in order to pass. Some common forms of assessment instruments include 'shadowing, cloze tests, written translation, memory tests of vocabulary, and interviews' (Lewis, 2004:90). The rigor of these translation tests and their high failure rates have led some to question the appropriateness of these assessment instruments which seem to expect the test takers to perform almost at the level of professional translators or interpreters while they are still taking training courses (Gentzler, 2001). The predictive power of the tests and the lack of objective assessment criteria have also been criticized by some who advocate research to correct these deficiencies (Arjona-Tseng, 1994; Rebenish, 2002).

These less crucial issues contrast starkly with the paucity of the discussion on the concepts of the central topics of test validity and reliability. The knowledge gap in these areas is so vast that we can do no more than sketch the problem here. Validity in interpreting and translation tests are tied up with problematic issues such as the nature of the competencies assessed, the models of learning underpinning the educational programs and the extent to which tests should reflect professional tasks. Reliability stands out as the priority problem that, according to Bachman (1991:163),

is concerned with 'how much of an individual's test performance is due to measurement errors or to factors other than the language ability we want to measure, and with the minimizing the effects of these factors on test scores'.

These factors can be grouped into 'test method facets', 'the attributes of the test taker that are not considered part of the language capabilities that we want to measure', and 'random factors that are largely unpredictable and temporary' (1991:164). Given that the latter two groups apply to the tests of any kind, we focus on 'test method facets' as criteria affecting the reliability of interpreting and translation assessment tests.

On the other hand, those EFL instructors fortunate enough to use multiple-choice and other brief response test item types have the luxury of measuring test reliability through *split-half* methods, where 'we divide the test into two halves and then determine the extent to which scores on these two halves are consistent with each other' (Bachman, 1991:172). The crucial requirement of split-half measures is that the performance on one half must be independent of the performance on the other half. Even if we could find ways to split the interpreting and translation tests (for example, odd versus even paragraphs, first half versus the second half, etc.) there is no way that the two halves can be independent; if they were, they would not constitute a text.

Indeed, the interpreting and translation tests seem to have much more in common with open-ended instruments like essays, where statistical methods of estimating reliability on the basis of individual test items are extremely difficult to apply. Highly applicable to the current issues and complications, the wider field of language education and assessment represents a solid source of knowledge to understand and improve the translation quality assessment practice. In retrospect, going back to 1956, Benjamin Bloom's Taxonomy of Educational Objectives: Cognitive Domain seems an appropriate hierarchy of fundamental human cognitive skills ordered from the lowest to the highest level(s) of complexity.

In 1956, Bloom attended the American Psychological Association in Boston to find possible solutions to the problems involved in organizing classified educational objectives. As a result, he reached a consensus with the APA to design his taxonomy to improve communication among examiners. In addition, the committee anticipated that the taxonomy would aid EFL teachers and test-takers through providing a basis for developing curricula, instructional techniques, and testing techniques.

- Knowledge: The lowest level in cognitive domain is the ability to simply recall the previously learned materials.
- Comprehension: A step beyond the knowledge level is the ability to grasp the meaning of the material.
- Application: A higher-level of understanding than the comprehension level is the ability to apply the learned material, methods, rules, concepts, laws, and theories to new and concrete situations.
- Analysis: It refers to the ability of breaking down material into its component so as to understand its internal structure.
- Synthesis: An ability to put parts together in a creative way in order to form a new whole is nearly the most difficult cognitive process.
- Evaluation: The most complex level in the cognitive hierarchy is concerned with the ability to appraise the value of received material.

Parks (2007) recorded that the major intention of Bloom's taxonomy is what the teachers expect students to know that can be arranged in a hierarchy from the least to the most complex.

Listed are some sample verbs for Bloom's six levels of cognitive domain.

Table 1. Bloom's Behavioral Objectives

Level	Behavioral Objectives
Knowledge	Write, List, Label, Name, State, Define
Comprehension	Explain, Summarize, Paraphrase, Describe, Illustrate
Application	Use, Compute, Solve, Demonstrate, Apply, Construct
Analysis	Analyze, Categorize, Compare, Contrast, Separate
Synthesis	Create, Design, Hypothesize, Invent, Develop
Evaluation	Judge, Recommend, Critique, Justify

Bloom's et al. (1956) believed that evaluating educational objectives is a constant concern in curriculum development. They claimed that their cognitive taxonomy could be utilized as a constructive set of guidelines for measuring every class of objectives from the simpler to the more complex.

2. Method

The participants in this study were 30 undergraduates majoring English translation at Islamic Azad University, Karaj Branch, Iran. They were selected based on the convenient method of sampling in the academic year of 2012. The participants were 25 to 39 years of age, of both genders and of similar nationality.

To homogenize the participants for their level of English proficiency, a Comprehensive English Language Test (CELT) was administered. The following week, a 60-minute, teacher-made English translation test, designed and validated by the researchers, was conducted.

The texts were graded for their difficulty levels as they became successively more difficult based on the six cognitive levels of Bloom's hierarchy of cognitive domain. For instance, to successfully translate the Application text, the participants had to *apply* their knowledge of English vocabulary to translate a text about charting techniques. Similarly, success in work with Evaluation text required the participants to *judge* the attitude of the author on the balanced budget proposal in America. The translated passages were holistically assessed by the researchers from a scale of 0 (the lowest quality) to 5 (the highest quality).

In order to determine the reliability of the placement test results, Cronbach's Alpha test was conducted for the total number of test items. The obtained Cronbach $\alpha = .606$, significant at p -value < 0.05 , demonstrated the relative consistency of the participants' performance on total test items. The researchers collected the data from the CELT and English Translation Test (ETT) both undertaken as a non-parametric Kruskal Wallis test of analysis of variance.

3. Results and Discussion

The scores obtained in the Comprehensive English Language Test (CELT) and English Translation Test (ETT) is summarized in Table 2.

Table 2. Descriptive Statistics for CELT and ETT

Test	K	Minimu m	Maximu m	Mean	Std. Deviation	Skewness	
						Statistic	Std. Error
CELT	72	24.00	56.00	41.18	1.64	9.28	-.662 .117
ETT	6	1.00	5.00	3.59	0.61	0.45	-.695 .141

As Table 2 represents, scores in the CELT and ETT have relatively similar ranges and averages, however, the standard deviation in the CELT reports a wider scope of scores than in ETT ratings. This diversity sounds justifiable as the participants were predisposed to several English comprehension courses prior to this research. Both sets of scores are negatively skewed.

To estimate the difficulty level of the texts, the researchers' assumption was in favor of an increasing difficulty from Knowledge to Evaluation texts, which could be demonstrated through the participants' quality of translation performance on the six English passages. They were carefully prepared based on the six levels of Bloom's Cognitive Domain (i.e., Knowledge, Comprehension, Application, Analysis, Synthesis and Evaluation) in his taxonomy of educational objectives. Finally, the researchers conducted a Kruskal Wallis non-parametric test to analyze the ratings.

Table 3. Kruskal Wallis Test for ETT Ratings

Texts		Sum of Squares	df	Mean Square	F	Sig.
KNOWLEDGE	Between Groups	12.719	21	.606	1.730*	.020
	Within Groups	3.500	10	.350		
	Total	16.219	31			
COMPREHENSION	Between Groups	22.135	21	1.054	1.723*	.030
	Within Groups	14.583	10	1.458		
	Total	36.719	31			
APPLICATION	Between Groups	44.792	21	2.133	3.506*	.020
	Within Groups	6.083	10	.608		
	Total	50.875	31			
ANALYSIS	Between Groups	33.958	21	1.617	1.481*	.040
	Within Groups	10.917	10	1.092		
	Total	44.875	31			
SYNTHESIS	Between Groups	31.667	21	1.508	1.552	.400
	Within Groups	14.333	10	1.433		
	Total	46.000	31			
EVALUATION	Between Groups	24.552	21	1.169	.784*	.000
	Within Groups	14.917	10	1.492		
	Total	39.469	31			

As Table 3 indicates, for Knowledge text, $F(21, 10)=1.730$ ratio was obtained which was significant at $p\text{-value}<.020$. For Comprehension text, the $F(21, 10)=1.723$ was obtained which was significant at $P<.030$. The F ratio was significant for Application text, too. ($F(21, 10)=3.506$, $P<.020$) Regarding Analysis text, the $F(21, 10)=1.481$ was significant at the $P<.040$. However, for Synthesis text, the $F(21, 10)=1.05$ turned out to be insignificant at $P<.400$. Finally, for Evaluation text, the $F(21, 10)=.784$ was significant at $P<.000$

To discuss the statistical results in this study, it is worth mentioning that the focus of this investigation was to (1) study the possible connection between the trainee translators' level of language proficiency and their translation performance, and to (2) assess the quality of their translation performance on the graded English texts based on the six levels of cognitive domain in Bloom's taxonomy of educational objectives. The researchers' initial assumption was that the trainee translators' English to Persian translation performance on these graded texts would show a descending pattern which would match the ascending difficulty of the texts.

1. The first research question was whether an English text demanding only the translators' knowledge of language stands at Level 1 of difficulty, according to Bloom's cognitive domain. Based on the findings in this study, the index of $F=1.730$ at $p\text{-value}<.05$ level of significance confirmed that it was? the easiest type of text for English to Persian translation.
2. The second research question was whether the text demanding only translators' comprehension ability stands at Level 2 of Bloom's cognitive domain. The findings supported this assumption as the F value decreased from 1.730 to 1.723 significantly at $p\text{-value}<.05$.
3. The third research question was whether translating an English text demanding only participants' application ability stands at Level 3 of Bloom's cognitive domain. Despite the value of F for Application text increased from .723 to 3.506, it was still significant at $p\text{-value}<.05$.
4. The fourth research question was whether a text only demanding the trainee translators' analytic ability stands at Level 4 of Bloom's cognitive domain. Statistically, the significant index of $F=1.481$ at $p\text{-value}<.05$ showed a decrease and supported its standing positing.
5. The fifth research question was whether a text only demanding the translators' synthetic ability stands at Level 5 of Bloom's cognitive domain. Unlike the other findings, the descending value of index $F=1.052$ was insignificant at $p\text{-value}<.400$. This finding rejected the stand of English Synthesis texts as more demanding for translation into Persian than Analysis texts and less difficult than Evaluation texts. In other words, statistically, the within-group differences outweighed the between-group differences which was incongruent to the researchers' initial assumption.
6. The sixth research question was whether translating an English text only demanding the translators' evaluating ability stands at Level 6 of Blooms' cognitive domain. The descending index of $F=.784$, significant at $p\text{-value}<.05$, confirmed the stand of Evaluation texts as the most difficult for English into Persian.

Interpreting the findings in the current study, the researchers concluded that Bloom's taxonomy of educational objectives was a breakthrough in grading English texts. Except for texts of Synthesis that were not statistically supported for an exact position on a scale of difficulty, the other five stands gave a clear-cut image for grading English texts. The researchers' explanation for the unexpected results mainly goes for the misinterpreting, narrowing down, or delimiting the number of behavioral objectives listed in Bloom's while selecting the synthesis text.

This research, however, provided the English translation trainers and test writers with a well-established psychological model of quality assessment through application of Bloom's 6 levels of cognitive domain which properly measured the English into Persian translation performance. The researchers strongly recommend further research on grading texts that more precisely operationalize Bloom's behavioral objectives underlying every level of cognitive domain. They believe that the more behavioral objectives encompass the selected texts, the more properly the text is graded for its level of difficulty. The length of the texts and the holistic or analytic quality assessment (of it) are questions, which require further experimentation. The researchers in this study focused on preparing texts with general topics, which evidently made the job easier for selecting participants and collecting data. Developing English for specific course contents (i.e., ESP) or selected texts is among the topics that may need further research.

Acknowledgements

The authors would like to extend their gratitude to Islamic Azad University, Karaj Branch for its financial supports to this study.

References

- Arjona-Tseng, E. (1994). A psychometric approach to the selection of translation and Interpreting students in Taiwan. In Lambert and Moser-Mercer (eds.) (1994), pp. 69-86.
- Bachman, L. F. (1991). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bell, R. (1991). *Translation and Translating: Theory and Practice*. London and New York: Longman.
- Bloom, B. S., Enghart, M. D., Furst, E. J., Hill, V. H., and Krathwohl, D. R. (1956). *Taxonomy of educational objectives*. Longmans Green: New York.

- Campbell, S. (1999). *The Translatability of a Christian Community*. New York: Fuller Theological Seminary.
- Campbell, S. & Hale, S. (1991). Translation and interpreting assessment in the context of educational measurement. In G. Anderman and G. Rogers (eds.) (1999), pp. 205-221.
- Faerch, C. & Kasper, G. (1987). *Strategies in Interlanguage Communication*. London: Longman.
- Gentzler, E. (2001). *Contemporary Translation Theories*. Clevedon: Multilingual Matters.
- Holmes, J. S. (1988b). *Translated Papers on Literary Translation and Translation Studies*. Amsterdam: Odopi.
- Holmes, J. S. (2004). The name and nature of translation studies. In L. Venuti (ed.) (2004), pp. 180-192.
- Krings, H. (1986). Translation problems and translation strategies of advanced German learners of French (L2). In J. House and S. Blum-Kulka (eds), pp. 262-76.
- Lambert, J. R. (1991). Shifts, opposition and goals in translation studies: towards a genealogy of concepts. In K. van Leuven-Zwart and T. Naaijken (eds), pp. 25-37.
- Lefevere, A. (1993). *Translating Literature: Practice and Theory in a Comparative Literature Context*. New York: The Modern Language Association of America.
- Lewis, P. (2004). The measure of translation effects. In L. Venuti (ed) (2004), pp. 256-75.
- Mason, I. (2003). Text parameters in translation: transitivity and institutional cultures. In L. Venuti (ed) (2004), pp. 470-81.
- Munday, J. (2001). *Introducing Translation Studies: Theories and Applications*. London and New York: Routledge.
- Munday, J. (2008). *Style and Ideology in Translation: Latin American Writing in English*. New York: Routledge.
- Parks, T. (2007). *Translating Style: The English Modernists and Their Italian Translations*. Manchester: St. Jerome.
- Pym, A. (2004a). Propositions in cross-cultural communication and translation, *Target*, 16 (1), pp.1-28.
- Rebenish, S. (2002). *Jerome*. London and New York: Routledge.

Appendix

Knowledge Text

The acronym UNICEF stands for United Nations International's Children's Emergency Fund. This organization was created by the United Nations in 1946 to provide emergency help to children in post-war Europe and China. UNICEF works with the government of individual countries to provide three kinds of services. It plans and develops programs, trains people to work in these programs, and provides the supplies and equipment to help the programs function. The programs that UNICEF develops cover five main areas: immunization, education for girls, child protection, early childhood development, and HIV/ AIDS.

Comprehension Text

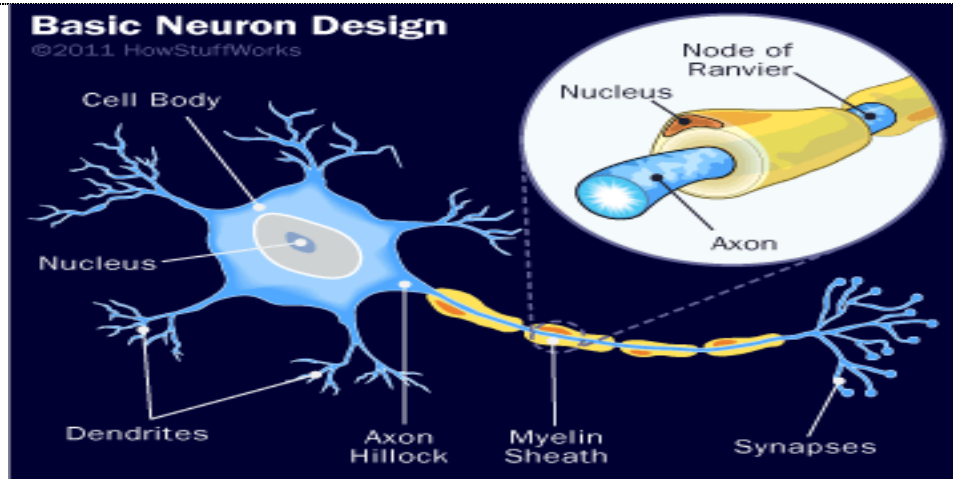
Since 1920 the population of China has doubled. With over one billion people today, China accounts for 23 percent of the world's population. This increase is the greatest obstacle for China's centrally controlled modernization plans. In order to solve this population problem, China began a 'One Child' policy. This meant that married couples had to limit their families to one child even if they would like to have more children. This policy has been most effective in the big cities where residents live in crowded apartments. In the countryside, however, the farmers say they are hard pressed to limit their families to one child since they need more children to help with farm work.

Application Text

Although every charting technique is useful in selling and demonstrating new methods to people, the multiple-activity process chart is the most popular. With the interrelated activities of several persons and machines, it is difficult to demonstrate new methods to people without this chart. Since the multiple-activity process chart clarifies unavoidable delay time for both workers and machines, it helps us to improve the utilization of these resources. This chart can specify the overall cycle time of a work system with multiple activities, which becomes the basis of the standard time.

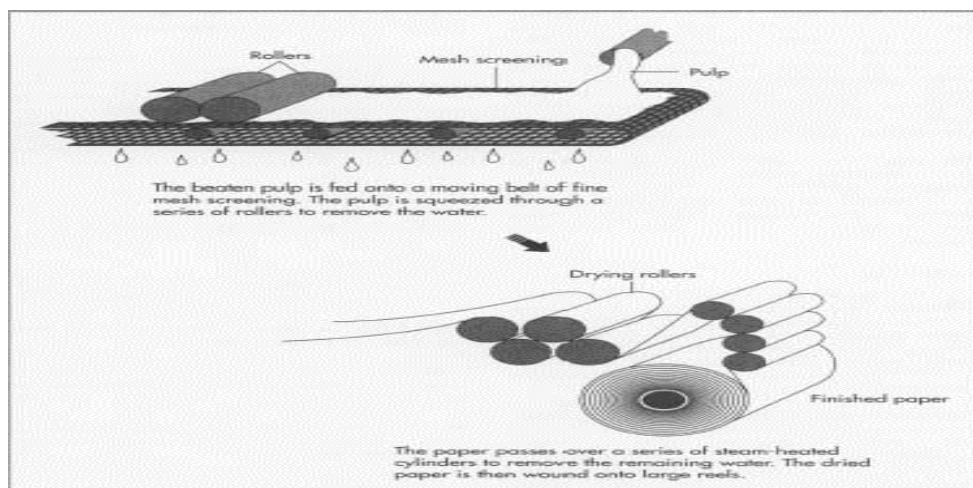
Analysis Text

Every animal you can think of: mammals, birds, reptiles, fish, amphibians has a brain. But the human brain is unique. Although it is not the largest, it gives people the power to speak, imagine and problem solve. It is truly an amazing organ. We will examine the structures of the brain and how each section controls our daily functions, including motor control, visual processing, auditory processing, sensation, learning, memory and emotions. Your brain is made of approximately 100 billion nerve cells, called neurons. Neurons have the amazing ability to gather and transmit electrochemical signals – think of them like the gates and wires in a computer. Neurons share the same characteristics and have the same makeup as other cells, but the electrochemical aspect lets them transmit signals over long distances (up to several feet or a few meters) and send messages to each other.



Synthesis Text

Several processes are commonly used to convert logs to wood pulp. In the mechanical process, logs are first tumbled in drums to remove the bark. The logs are then sent to grinders, which break the wood down into pulp by pressing it between huge revolving slabs. The pulp is filtered to remove foreign objects. In the chemical process, wood chips from de-barked logs are cooked in a chemical solution. This is done using a solution of sodium hydroxide and sodium sulfide that results in chips dissolving into a pulp. Next the pulp is sent through filters. Bleach may be added at this stage, or colorings. The pulp is then sent to the paper plant for final processing.



Evaluation Text

A continual aspect of the budget is the deficit. Franklin Roosevelt believed that the government should spend more than it took in, in order to get the economy running again. He thought this would be a temporary adjustment to the budget, but it remained unchanged. When the Balanced Budget Act was proposed, there were a lot of people who protested it. It was thought that it would not work because if the budget were balanced the economy would slow. The government would have to cut many government funded programs, and this would cause more problems with the budget and the overall economic outlook.