The Development of Scoring Criteria for a New Picture Naming Task

Ferzin MAHAVA* [1], Christine SHEPPARD [2], Laura MONETTA [3,4], Vanessa TALER [2,5]

[1] Student, University of Ottawa, Canada
[2] Researcher, Bruyère Research Institute, Canada
[3] Researcher, Université Laval, Canada
[4] Researcher, Centre de Recherche de l'Institut Universitaire en Santé Mentale de Québec, Canada
[5] Professor, University of Ottawa, Canada

** Auteur(e) correspondant | Corresponding author :* ferzin.mahava@gmail.com

Résumé :

(traduction)

Objectif : Le but de cette étude est de développer un système de notation pour une nouvelle tâche de dénomination afin d'évaluer la performance de dénomination de jeunes adultes (18 à 30 ans) et d'ainés (65+ ans) unilingues anglophones, unilingues francophones, et bilingues anglais-français. Cette nouvelle tâche de dénomination servira de service de santé important pour aider à diagnostiquer et évaluer les personnes âgées souffrant de troubles cognitifs, tout en servant d'outil pédagogique pour les fournisseurs de soins de santé.

Matériels et méthodes : Cent-vingt images sur fond blanc dans PowerPoint ont été présentées aux participants, suivant le même ordre aléatoire. La tâche de dénomination consistait à nommer l'image affichée à l'ordinateur. Les participants unilingues ont passé le test dans leur langue maternelle et ceux bilingues ont passé le test en anglais, en français, et dans les deux langues. Un critère de notation a été établi selon les réponses apportées.

Résultats : Des critères de notation stricts et souples ont été développés pour la tâche de dénomination. Huit images ont été supprimées de la tâche de dénomination originale en raison de leur faible qualité et/ou clarté, de l'incapacité des participants à les nommer, ou de la grande variété de réponses qui leur étaient associées. Selon les critères stricts et souples, la performance de dénomination des jeunes adultes et des ainés unilingues francophones et anglophones était similaire. Les personnes bilingues ont le mieux réussi lors du test bilingue et le moins bien réussi en français. Parmi tous les groupes d'âge et de langue, les résultats des participants bilingues en français étaient les plus faibles.

Conclusion : La tâche de dénomination semble appropriée pour les participants unilingues francophones et anglophones. Les résultats suggèrent qu'un test bilingue devrait être utilisé pour évaluer les personnes bilingues anglais-français.

Mots-clés : Dénomination d'image, bilinguisme, vieillissement, tâche de dénomination

Abstract:

Objective: The purpose of the study was to develop a scoring system for a novel naming task suitable for assessing naming performance in younger (18-30 years) and older (65+ years) adults in monolingual English, monolingual French, and English-French bilingual groups. This novel naming task will serve as an important health service to help diagnose and assess cognitively impaired older individuals, while also serving as an educational tool for healthcare providers.

Materials and Methods: The Naming Task consists of 120 images organized in the same randomized order, and are shown on a white background displayed on a computer screen using PowerPoint. Participants are instructed to name the image displayed. Monolinguals completed the test in their native language and bilinguals completed the test in English only, French only, and a bilingual administration. Scoring criteria was established based on the responses from testing.

Results: Strict and lenient scoring criteria developed for the Naming Task are presented. Eight items were removed from the original Naming Task due to quality and/or clarity, inability to name the image, or too many alternate responses. Performance in monolingual English and French was similar in younger and older adults for strict and lenient scoring. Bilinguals performed better with bilingual administration and worse with French administration, where scores were the lowest of all age and language groups.

Conclusion: The Naming Task appears to be suitable for monolingual French and English individuals. Results suggest that a bilingual administration should be used when testing English-French bilinguals.

Keywords:

Picture naming, bilingualism, aging, naming task

Introduction

Despite the overwhelming increase of bilingualism in Canada, there are no appropriate tools to assess language abilities in older English-French bilingual speakers. A new Naming Task will serve as a tool for healthcare providers to assess naming abilities in bilingual adults. This may be important when assessing older adults for medical conditions that impact language abilities, such as dementia and aphasia. The purpose of the present study is to develop a scoring system for a novel naming task that is suitable for assessing naming performance in monolingual English, monolingual French, and English-French bilinguals. Upon scoring criteria development, this novel naming task will serve as an important health service to help diagnose and assess cognitively impaired older individuals.

Two types of scoring criteria were developed for the Naming Task: strict and lenient scores. Strict scores represented the formal name for an item, while lenient scores included acceptable synonyms or slang terms. The analysis presented in this paper will determine which names are used the most often for each item and establish a clear set of guidelines for strict and lenient scoring in both English and French. Performance across groups will be compared on the strict and lenient scoring criteria, in order to examine the impact of language administration on bilingual performance and to determine if the test is suitable for all language groups.

Literature Review

In the recent decade, research has begun exploring the impact of bilingualism on cognition, especially in the areas of executive function and language. This research has demonstrated that, relative to monolinguals, bilingual individuals show superior performance on tasks of executive function (e.g., inhibition of task-irrelevant information; Adesope, Lavin, Thompson, & Ungerleider, 2010; Bialystok, 2009; Bialystok, Craik, Green, & Gollan, 2009), but poorer performance on language tasks (e.g., picture naming tasks) (Gollan, Montoya, Fennema-Notestine, & Morris, 2005; Roberts, Garcia, Desrochers, & Hernandez, 2002). In addition, bilingualism can be seen as a protective factor, as research with an immigrant sample living in Toronto has suggested that bilingualism may delay the onset of dementia by five years in older adults (Bialystok, Craik, & Freedman, 2007; Craik, Bialystok, & Freedman, 2010).

The Boston Naming Test (BNT) is a widely used clinical picture-naming task, where patients are asked to name the image displayed (Kaplan, Goodglass, & Weintraub, 1983). Overall, individuals show a decline in naming ability as they age (Kaplan et al., 1983), specifically after the age of 70 (Brouillette et al., 2011). Research examining the utility of the BNT with bilinguals has shown that monolinguals tend to outperform bilinguals and the level of difficulty for the test likely differs between languages (Roberts et al., 2002). For example, in a study comparing English-speaking monolinguals, bilingual Spanish-English speakers, and bilingual English-French speakers, both bilingual groups scored significantly worse than the monolingual English participants (Roberts et al., 2002). Furthermore, bilinguals have demonstrated difficulty with verbal fluency, frequent tip-of-the-tongue states, and longer picture naming latencies (Bialystok, 2009), even when completing the task in their dominant language (Gollan & Acenas, 2004). Additional studies have indicated that bilinguals perform worse on naming tasks such as the BNT, both in measures of accuracy (Bialystok, Craik, & Luk, 2008; Kohnert, Hernandez, & Bates, 1998) and response time (Gollan et al., 2005; Gollan, Fennema-Notestine, Montoya, & Jernigan, 2007; Ivanova & Costa, 2008; Roberts et al., 2002).

Research with French Canadians suggests that the French translation of the BNT does not account for cultural appropriateness, which is important when administering the test in a language other than the one in which it was originally developed (Roberts & Doucet, 2011). Specifically, research suggests that the French translation of the BNT is not acceptable for assessing naming abilities in English-French bilinguals or in monolingual French individuals (Roberts & Doucet, 2011; Sheppard, Kousaie, Monetta, & Taler, 2016). It has been suggested that when there is a large inconsistency in naming certain items, these items should be removed or the items should be changed in their order of difficulty (Roberts & Doucet, 2011). For example, research with older adults from Quebec City indicated that there were 13 BNT items with multiple acceptable synonyms (e.g., “seahorse” can either be “hippocampe” or “cheval de mer”) and an additional six items that had no clear acceptable response (e.g., “globe”), as native speakers in French disagree on the name of the item (Roberts & Doucet, 2011). Additional research comparing monolingual English and French speakers to English-French bilinguals on the BNT demonstrated that a French administration of the task consistently yielded poorer scores, even in the French monolingual group (Sheppard et al., 2016). Furthermore, after matching for underlying naming ability, differential item functioning analyses suggested that a significant number of items functioned differently across the three participant groups and in different languages of administration (Sheppard et al., 2016), suggesting that the BNT is not equivalent in English and French.

Materials and Methods

Participants

Six groups of participants were included in this study: younger (n = 44) and older (n = 64) monolingual-English speakers, younger (n=30) and older (n = 30) monolingual-French speakers, and younger (n = 48) and older (n = 52) bilingual English-French speakers. Young adults were aged 18 to 30 and older adults were aged 65 or older. Monolingual English participants and bilingual English-French participants were recruited and tested in the Ottawa-Gatineau region, while monolingual French speakers were recruited and tested in Quebec City. Younger adults were recruited through word of mouth and local undergraduate populations, while older adults were recruited through advertisements in community centres, grocery stores, and newspapers. Monolingual participants had either limited or no exposure to languages other than their native language. Bilinguals had limited exposure to languages other than French and English. All bilingual participants were proficient in both English and French before the age of 13 and self-reported their proficiency in French and English using a 5-point Likert scale (see Table 1) on measures of auditory comprehension, reading, speaking, and writing.

Naming Task

The Naming Task consists of 120 images, 100 of which were selected from the coloured Snodgrass set (Rossion & Pourtous, 2004) and the remaining 20 were developed by Dr. Taler, the lead researcher in this study. The Snodgrass images were selected based on their array of difficulty and strong name agreement, while the additional images were created based on the same colour scheme as the Snodgrass set, but with a higher level of naming difficulty. The images were organized in the same randomized order for all participants and were shown on a white background displayed on a computer screen using PowerPoint. Participants were instructed to identify the image on the screen and the research assistant was instructed to record all answers given by the participant.

Neuropsychological Battery

Participants completed a neuropsychological battery, including the forward and backward digit span subtests of the Wechsler Adult Intelligent Scale-Third Edition (Wechsler, 1997); the Montreal Cognitive Assessment (Nasreddine et al., 2005); a version of the Stroop colour-word interference test (Stroop, 1935) in which the number of items produced in 45 seconds was recorded in each of the three conditions (word reading, color naming, and incongruent colour naming); the 64-item Wisconsin Card Sorting Test (Grant & Berg, 1948); and category (animal) and letter (FAS) verbal fluencies (Benton & Hamsher, 1976). Monolingual participants completed the verbal fluency tasks in their native language and bilingual participants completed the tasks in English, in French, and in an administration where they could respond in either language. The neuropsychological battery was administered to demonstrate that all study participants had normal cognitive function. See Table 2 for demographics and neuropsychological performance across all groups.

Table 1 Mean calculation ± standard deviation of proficiency by modality for both English and French for bilingual younger (n = 48) and bilingual older (n = 52) participants. Ranking followed a 5-point Likert scale (1 = no ability; 5 = native-like ability).

	Young Adults				Older Adults			
	English		French		English		French	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Auditory Comprehension	5.00	0.00	4.73	0.51	4.94	0.24	4.78	0.50
Reading	4.95	0.22	4.64	0.53	4.92	0.27	4.76	0.43
Speaking	4.93	0.27	4.49	0.59	4.94	0.34	4.67	0.52
Writing	4.68	0.36	4.23	0.83	4.83	0.43	4.51	0.78

Table 2

Demographic and neuropsychological performance by participant group (mean ± standard deviation). Verbal fluency scores for bilingual groups are reported where participants could answer in either language. MoCA = Montreal Cognitive Assessment; Digit Span= Wechsler Adult Intelligent Scale-Third Edition; WCST = 64-item Wisconsin Card Sorting Test; FAS = letter verbal fluencies; Animals = category verbal fluencies.

	Younger Adults						Older Adults					
	Monolingual English (n = 44)		Monolingual French (n = 30)		Bilingual (n = 48)		Monolingual English (n = 64)		Monolingual French (n = 30)		Bilingual (n = 52)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age (years)	21.84	1.84	21.80	2.47	21.10	2.00	70.81	2.73	72.60	6.59	72.23	6.36
Education (years)	15.70	1.21	15.13	1.38	15.71	1.52	15.23	2.70	16.20	2.57	16.38	2.64
MoCA (/30)	28.32	1.38	27.53	2.64	27.83	1.51	27.74	1.52	27.50	1.36	27.71	1.73
Digit Span												
Forwards (/16)	11.25	2.09	12.53	2.29	10.95	2.42	10.79	1.89	10.47	2.40	10.55	2.03
Reverse (/14)	6.95	1.95	8.77	2.62	7.80	2.23	7.56	2.19	7.07	2.03	7.92	2.38
WCST (categories /6)	4.28	1.01	4.70	0.47	4.59	0.64	3.61	1.21	4.07	0.69	3.62	1.23
Stroop												
Word	104.71	17.38	115.23	15.86	108.95	14.37	95.08	14.87	104.07	16.05	96.09	15.34
Colour	78.29	13.84	86.77	14.37	76.93	10.96	66.16	12.83	73.10	14.07	62.17	12.44
Inference	52.69	11.70	48.67	8.26	52.78	8.44	34.46	7.90	32.07	9.85	36.68	8.25
Verbal Influency												
FAS	40.59	13.10	38.23	8.24	38.90	10.79	40.88	14.03	36.37	10.13	38.88	15.75
Animals	24.02	5.44	22.67	5.26	22.68	6.69	21.23	4.67	18.27	4.62	19.53	6.79

Procedure

All monolingual participants completed the testing in one session of two hours, while bilingual participants completed the testing in two sessions of two hours each. All bilingual participants completed the Naming Task in three administrations: English only, French only, and either-language where they could respond in either English or French. Two language administrations were completed in the first testing session, while the third administration was completed in the second testing session.

The study procedures adhered to federal guidelines for protection of human research participants and received ethical approval from the Research Ethics Board at the Bruyère Research Institute, Laval University, and the University of Ottawa. Participants were remunerated $10/ hour for all testing completed and provided informed consent prior to participating.

Development of Scoring Criteria

Dr. Taler developed preliminary scoring criteria for the Naming Task in English and French; these scoring criteria formed the basis of the strict and lenient scoring protocol that was developed for this study. First, the data from each participant were scored based on the preliminary scoring criteria, wherein one point was awarded for each correct answer. Percentages were then calculated for each image based on the number of participants who named the image correctly. During this process, alternative answers provided by participants were recorded. Two independent reviewers went through each item to determine the strict and lenient scoring criteria. The strict scoring criteria were selected based on the most frequent response provided by participants (i.e., a minimum of 50%) and/or the most formal or known name used in society. Lenient responses were selected based on synonyms (e.g., “ironing board” vs. “ironing table”), clarity of the image (e.g., “violin” vs. “viola”), culturally relevant slang terms (e.g., “baby carriage” vs. “pram”), and shortened names for the image (e.g., “green pepper” vs. “pepper”). The two independent researchers then met to discuss their findings. Discrepancies were resolved through discussion and all established scoring criteria were verified by three additional researchers. See Appendix A for a list of strict and lenient responses for each item.

Results

Items Recommended for Removal

Eight items were recommended for removal in English and French: stirrup, gavel, beetle, barn, blouse, and flute were removed due to the clarity and/or quality of the image; rickshaw was removed because no younger or older monolingual French participants could name the image; and necklace was removed as there were too many alternative names for these image (e.g., for necklace: “pearls”, “string of pearls”, “pearl necklace”, and “necklace”).

Overall Task Performance

Figures 1 and 2 present an overall summary of task performance by age and language group according to strict and lenient scoring criteria. The largest difference in naming abilities between older and younger adults is seen in the bilingual French administration groups. Overall, older adults performed better than younger adults in all language categories. The only group where younger participants scored higher than older participants was the monolingual French group, and younger participants scored an average of one item higher (strict and lenient).

For both younger and older adult groups, monolingual English participants had the highest overall score across the task, ranging from an average of 99 correct items using strict scoring and 106 correct items using lenient scoring, out of 120 items. Bilingual English-French participants were able to correctly name an average of 92 and 94 (strict and lenient scoring, respectively) of the items when completing the test in English; however, this increased to 95 and 102 (strict and lenient scoring, respectively) when responses were accepted in either language. The majority of bilingual participants in the bilingual administration responded in English (i.e., 52% of older adults and 62% of younger adults). The average number of items named correctly did not improve by more than five items in any group when lenient scoring was added.

Results by Item

Table 3 represents the percentage of participants who correctly identified each item under strict and lenient scoring.

Figure 1 Average number of images named under strict scoring criteria by age and language group.

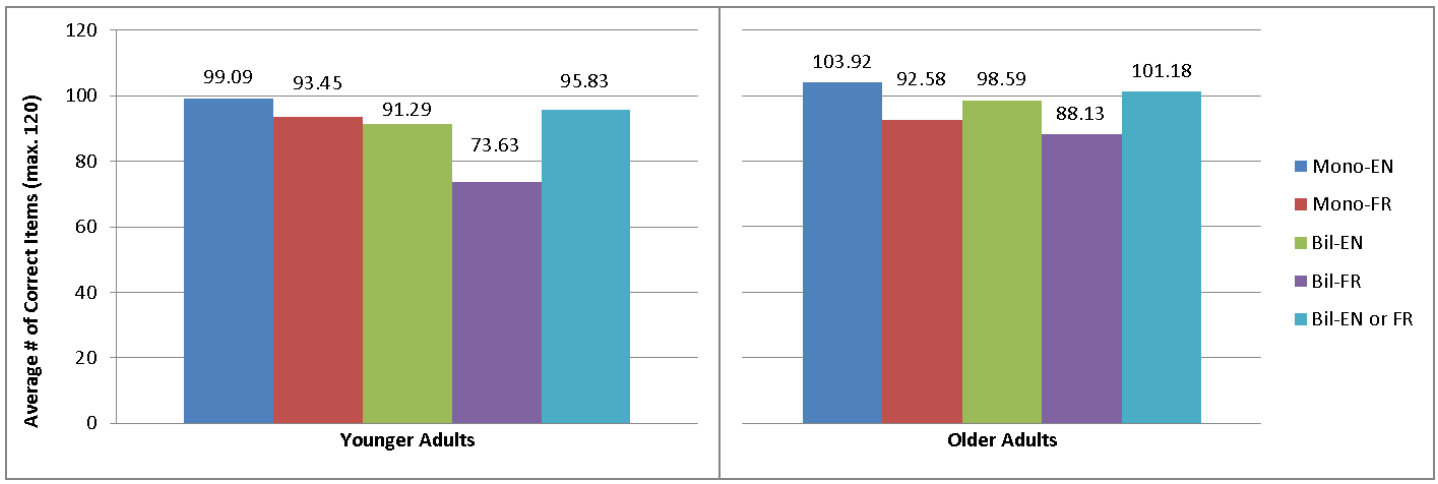

Figure 2 Average number of images named under lenient scoring criteria by age and language group.

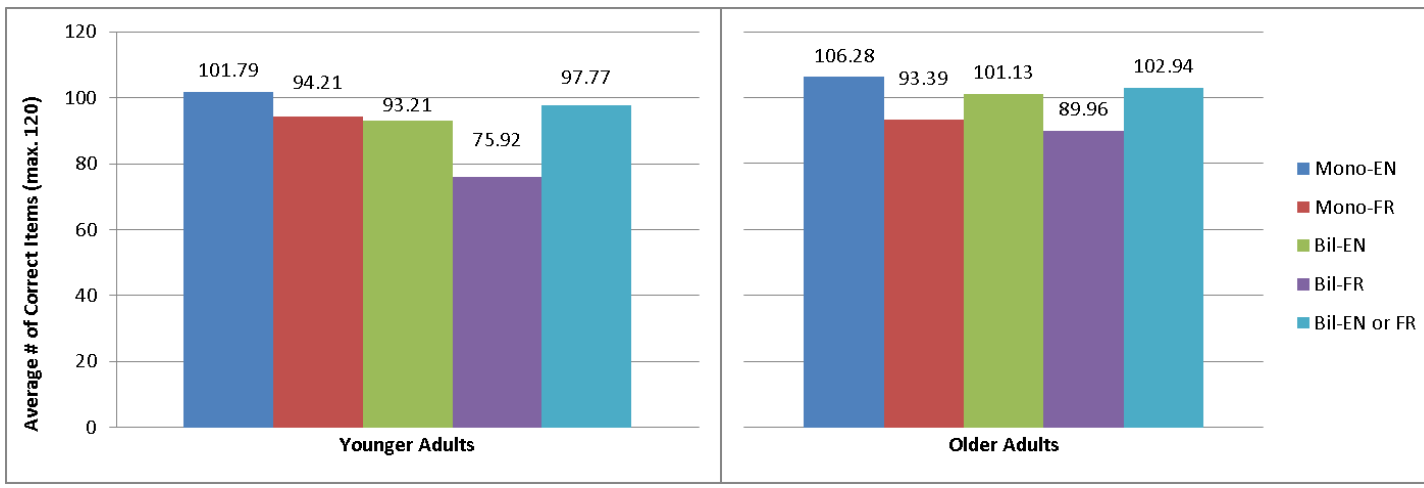

Table 3 Percentage of correct item responses for strict and lenient scoring for participants in monolingual and bilingual groups. ME = Monolingual English; MF = Monolingual French; YA = Younger adults; OA = Older adults; St = Strict; Len = Lenient; Eng = English Administration; Fre = French Administration; Bil = Bilingual Administration.

Item	**ME** (%)				**MF** (%)				**Bilingual YA** (%)						**Bilingual OA** (%)					
	YA		**OA**		**YA**		**OA**		**Eng**		**Fre**		**Bil**		**Eng**		**Fre**		**Bil**	
	St	**Len**	**St**	**Len**	**St**	**Len**	**St**	**Len**	**St**	**Len**	**St**	**Len**	**St**	**Len**	**St**	**Len**	**St**	**Len**	**St**	**Len**
crown	100	100	100	100	100	100	100	100	98	98	94	94	98	98	100	100	98	98	100	100
helicopter	100	100	100	100	100	100	100	100	100	100	100	100	100	100	98	98	100	100	98	98
barrel	98	98	98	98	55	55	71	71	96	96	75	75	100	100	100	100	94	94	100	100
tiger	100	100	95	95	100	100	87	87	100	100	98	98	100	100	100	100	96	96	100	100
rolling pin	88	88	97	97	97	97	100	100	54	54	48	48	81	81	93	93	77	77	96	96
spool of thread, spool	51	72	80	91	86	69	65	71	15	44	38	52	47	62	69	89	42	60	80	86
violin, fiddle	100	100	98	98	90	90	94	94	98	98	94	94	98	98	94	94	94	94	98	98
iron	100	100	100	100	97	97	100	100	88	88	65	65	96	96	98	98	96	96	100	100
alligator, crocodile	100	100	100	100	100	100	94	94	100	100	100	100	100	100	100	100	96	96	100	100
pliers	72	72	95	95	100	100	100	100	56	56	65	65	77	77	91	91	90	90	98	98
kangaroo	100	100	100	100	100	100	90	90	100	100	98	98	100	100	98	98	98	98	100	100
duck	95	95	91	91	97	97	87	87	98	98	96	96	100	100	96	96	94	94	92	92
guitar	100	100	95	95	100	100	94	94	100	100	100	100	100	100	98	98	94	94	96	96
trombone	77	77	72	72	66	66	45	45	75	75	71	71	72	72	80	80	75	75	78	78
well	93	95	100	100	100	100	97	97	94	94	65	65	98	100	98	98	92	92	96	96
rhinoceros	93	100	92	92	100	100	81	81	100	100	96	96	100	100	89	93	94	94	90	90
basket	98	98	98	98	100	100	100	100	100	100	88	88	98	98	98	98	98	98	100	100
lobster	95	95	100	100	97	97	100	100	90	90	60	60	89	89	96	96	77	77	98	98
cummer-bund	30	30	83	83	14	14	0	0	19	19	4	4	26	26	63	63	13	13	66	66
pipe	100	100	100	100	100	100	100	100	96	96	92	92	100	100	100	100	100	100	100	100
belt	100	100	97	97	100	100	100	100	100	100	98	98	100	100	98	98	96	96	100	100
ostrich	74	84	89	89	76	76	58	68	83	83	71	71	85	85	83	87	75	75	84	86
ottoman; hassock	60	77	42	84	59	66	55	55	27	33	23	27	45	49	41	67	40	40	54	64
chest of drawers; dresser	77	77	91	91	66	66	65	65	56	56	21	21	62	62	83	83	33	33	68	68
ruler	100	100	100	100	100	100	100	100	98	98	96	96	96	96	100	100	98	98	100	100
spinning wheel	30	30	92	92	14	14	77	77	19	19	4	4	23	23	76	76	38	38	80	80
asparagus	93	93	98	98	90	90	97	97	94	94	77	77	94	94	96	96	90	90	100	100
candelabra	21	30	78	88	0	10	26	26	17	38	13	23	15	34	72	72	38	42	62	66
leopard; cheetah	91	95	77	78	69	76	61	61	106	75	81	85	77	85	78	81	65	65	78	82
racquet	91	91	100	100	100	100	100	100	85	92	94	94	98	98	96	96	92	94	96	100
sheep	84	84	97	97	93	93	100	100	85	85	92	92	94	94	83	83	79	79	94	94
door knob	93	95	92	92	100	100	97	97	79	79	60	60	81	81	83	83	94	94	94	94

Item	ME (%)				MF (%)				Bilingual YA (%)						Bilingual OA (%)					
	YA		OA		YA		OA		Eng		Fre		Bil		Eng		Fre		Bil	
	St	Len	St	Len	St	Len	St	Len	St	Len	St	Len	St	Len	St	Len	St	Len	St	Len
ear	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
boot	100	100	97	97	97	97	97	97	100	100	100	100	100	100	100	100	92	92	98	98
ring	100	100	98	98	97	97	90	90	100	100	83	83	100	100	98	98	90	90	98	98
grass hopper	79	98	97	97	90	90	74	74	83	92	63	63	85	96	83	93	79	79	86	90
nail file	77	84	88	88	66	66	65	65	81	83	52	52	85	85	85	91	63	63	90	92
screwdriver	95	95	100	100	100	100	100	100	92	92	77	77	98	98	100	100	92	92	100	100
glasses	98	98	95	95	100	100	100	100	100	100	98	98	100	100	98	98	100	100	100	100
record player	81	88	63	78	59	59	90	90	63	75	15	15	66	74	44	70	56	56	70	80
anchor	98	98	98	98	90	90	97	97	96	96	58	58	96	96	98	98	92	92	98	98
pineapple	100	100	97	97	100	100	97	97	100	100	98	98	100	100	98	98	90	90	100	100
nut	63	63	66	66	28	28	48	48	44	44	6	6	38	38	46	46	27	27	50	50
bridle	33	33	48	48	7	7	3	3	21	21	4	4	19	19	35	35	27	27	34	34
hanger	100	100	89	89	97	97	100	100	88	90	42	42	89	89	93	93	65	65	92	92
hammer	100	100	100	100	100	100	100	100	100	100	88	88	98	98	98	98	100	100	100	100
abacus	49	49	84	84	45	45	87	87	42	42	2	2	47	47	74	74	33	33	78	78
eagle	84	95	84	100	93	93	58	65	90	94	81	81	94	100	81	91	87	88	90	94
artichoke	40	40	70	70	34	34	81	81	33	33	23	23	34	34	69	69	58	58	52	52
lightswitch	98	98	95	98	62	62	42	42	92	94	29	29	96	96	78	91	29	29	78	84
carafe	16	16	55	55	34	34	26	26	10	10	10	10	17	17	50	50	50	50	44	44
eye	98	98	100	100	100	100	97	97	100	100	98	98	100	100	96	96	100	100	100	100
mushroom	100	100	97	97	100	100	100	100	98	98	94	94	100	100	96	96	98	98	98	98
ironing board	98	100	100	100	72	97	81	100	81	85	42	52	87	91	94	100	85	94	98	98
wrench	100	100	92	92	52	52	48	48	83	81	19	19	72	72	74	74	37	37	70	70
onion	100	100	97	97	97	97	100	100	98	98	98	98	98	98	98	98	98	98	98	98
centaur	56	56	33	33	62	62	19	19	38	38	25	25	45	45	31	31	35	35	32	32
axe	95	95	94	94	100	100	100	100	100	98	67	67	98	98	96	96	90	90	98	98
nail	100	100	98	98	86	86	100	100	88	90	85	85	98	98	98	98	98	98	100	100
squirrel	100	100	100	100	100	100	100	100	98	98	98	98	98	98	100	100	98	98	100	100
lips, mouth	100	100	100	100	100	100	87	87	100	100	98	98	100	100	98	98	100	100	100	100
mitten	81	81	64	89	93	93	100	100	77	81	90	90	89	91	80	87	94	94	92	100
cannon	93	93	89	89	100	100	90	90	96	96	92	92	98	98	96	96	96	96	96	96
stroller, carriage	70	70	11	11	90	90	84	84	56	56	27	27	55	55	28	28	21	21	28	28
gorilla	88	88	83	83	97	97	39	39	88	88	83	83	89	89	70	70	79	79	74	74
pomegran-ate	91	91	80	80	79	79	55	55	83	83	46	46	87	87	81	85	46	48	78	80
wagon	93	93	92	92	24	24	29	29	56	56	42	42	64	64	89	89	52	52	80	80
tambourine	91	91	66	66	48	48	19	19	67	67	54	54	64	64	54	54	54	54	62	62
heart	100	100	100	100	100	100	97	100	100	100	100	100	100	100	98	98	100	100	100	100
zebra	100	100	100	100	100	100	94	94	98	98	98	98	100	100	100	100	92	92	100	100
screw	100	100	98	98	100	100	100	100	96	96	75	75	98	98	98	98	96	96	98	98
celery	74	74	91	91	76	76	87	87	77	77	77	77	77	77	93	93	92	92	90	90

Item	ME (%)				MF (%)				Bilingual YA (%)						Bilingual OA (%)					
	YA		OA		YA		OA		Eng		Fre		Bil		Eng		Fre		Bil	
	St	Len	St	Len	St	Len	St	Len	St	Len	St	Len	St	Len	St	Len	St	Len	St	Len
calipers	7	7	36	36	0	0	3	3	8	8	2	2	9	9	19	19	12	12	14	14
stool	98	98	100	100	72	72	71	71	88	88	25	25	81	81	89	89	54	54	88	88
seahorse	93	93	91	91	100	100	65	65	94	94	35	56	89	94	78	78	31	60	80	82
bow	88	98	95	100	100	100	94	100	90	96	19	77	83	89	93	96	23	92	74	98
rollerskate	93	93	92	92	48	48	100	100	65	65	73	73	81	81	85	85	92	92	92	92
glove	100	100	98	98	97	97	100	100	96	96	92	92	98	98	96	96	100	100	100	100
peacock	86	86	95	95	86	86	94	94	73	73	31	31	89	89	69	69	67	67	90	90
vest	100	100	91	91	45	45	90	90	98	98	88	88	96	96	100	100	88	88	96	96
kettle	91	91	100	100	97	97	100	100	83	83	58	58	85	85	93	93	92	92	96	96
Bunsen burner	53	53	70	70	14	14	13	13	52	52	42	42	64	64	56	56	21	21	50	50
colander	37	98	72	83	90	90	90	90	19	71	29	29	40	79	44	69	42	42	66	78
coat	67	100	48	92	90	90	81	81	71	100	94	94	72	100	57	96	88	88	64	92
trumpet	91	91	70	70	100	100	55	55	92	92	90	90	91	91	76	76	85	85	72	72
trowel	28	28	59	59	0	7	0	10	8	8	2	15	13	13	35	35	12	17	30	30
raccoon	100	100	86	86	100	100	52	52	94	94	77	77	100	100	83	83	67	67	84	84
salt shaker	91	100	89	89	66	86	97	100	83	96	38	88	89	98	91	94	83	87	98	100
arrow	93	93	100	100	100	100	94	94	96	96	83	83	98	98	100	100	96	96	94	94
accordion	93	93	95	95	76	76	100	100	85	85	83	83	89	89	93	93	96	96	98	98
green pepper, bell pepper	98	98	98	98	100	100	100	100	96	96	42	77	89	94	94	94	58	90	82	100
broom	98	98	100	100	100	100	100	100	98	98	85	85	100	100	100	100	94	94	100	100
top	70	77	97	97	97	97	97	97	50	50	65	65	91	91	72	72	79	79	90	90
pitcher, jug	88	88	98	98	55	55	26	26	65	65	27	27	62	62	89	89	42	42	76	76
chisel	47	47	86	86	17	17	23	23	44	44	6	6	43	43	69	69	33	33	60	60
metronome	49	49	81	81	45	45	61	61	52	52	50	50	60	60	76	76	75	75	84	84
sled	91	98	84	86	86	86	87	87	77	88	38	38	85	96	72	76	58	58	82	82
hand	100	100	100	100	100	100	100	100	96	96	98	98	98	98	98	98	100	100	98	98
monocle	51	51	83	83	31	31	52	52	40	40	17	17	45	45	70	70	69	69	72	72
thimble	84	84	91	91	59	59	97	97	58	58	15	15	64	64	80	80	67	67	88	88
corn	100	100	100	100	93	93	74	74	100	100	96	96	98	98	100	100	96	96	94	94
clothespin, clothes peg	77	77	100	100	97	97	100	100	50	50	50	50	77	77	89	89	77	77	88	88
chicken	70	88	61	94	69	69	94	94	81	83	81	81	85	87	61	91	90	90	84	90
harp	100	100	97	97	90	90	94	94	90	90	75	75	96	96	98	98	92	92	98	98
pumpkin	100	100	98	98	100	100	97	97	98	98	98	98	100	100	93	93	96	96	100	100
watering can	74	77	97	97	90	90	100	100	35	35	48	48	72	72	81	81	83	83	86	88
saw	100	100	100	100	100	100	100	100	94	94	83	83	96	96	100	100	98	98	100	100
dragonfly	98	98	77	77	90	90	68	68	81	81	56	56	98	98	48	48	50	50	70	70
pear	98	98	100	100	100	100	100	100	98	98	77	77	100	100	100	100	98	98	100	100
rocking chair	91	91	100	100	93	93	94	94	92	92	63	63	94	94	96	96	85	85	94	94
windmill	98	98	98	98	97	97	100	100	83	83	58	58	96	96	94	94	85	85	98	98
butterfly	100	100	100	100	100	100	100	100	100	100	96	96	98	98	98	98	90	90	100	100

Analysis 1: Strict and Lenient Scoring Differences. There were a number of items where performance improved by one to five extra items once lenient criteria was taken into consideration. The following is a list of items where percentages improved once lenient scoring was included, in both English and French for all language groups: spool of thread, ottoman, candelabra, leopard, eagle, ironing board, bow, coat, and salt shaker. Additionally, there were a number of items that scored higher once lenient scores were included in English only: grasshopper, record player, beetle, light-switch, mitten, colander, and sled; and in French only: hippocampe, truelle, and poivron.

Analysis 2: Language Group Differences. Bilingual participants performed more poorly on the task than monolingual participants in their respective languages. The difference was most extreme when comparing the monolingual French participants and the bilingual-French administration. While there was a similar pattern of results shown with the monolingual English participants and the bilingual-English administration, the performance differences were not as great (i.e., smaller difference between groups) or consistent (i.e., not as many items displaying group differences). It should be noted that there are a small number of items where bilingual English-French speakers scored better than the monolingual groups. In English, these items include cannon, celery, and flute. In French, these items include cyclo-pousse, lèvres, wagon, and bec Bunsen.

Analysis 3: Age Differences. The following is a list of items that had large generational differences, where younger adults scored higher than older adults: necklace, centaur, stroller, gorilla, tambourine, trumpet, and racoon. However, overall, older adults scored higher than younger adults in all languages and language administration groups.

Discussion

The purpose of this study was to develop scoring criteria for a new bilingual naming task, as it will serve as an important health service for cognitively impaired older adults. Older and younger participants were tested using a preliminary scoring criteria to determine if the test was appropriate for both English- and French-speaking individuals. Although the task can easily be administered to all groups, there are differences in how each group of participants performs based on their age group, language group, and for the bilingual participants, language of administration.

Allowing lenient scoring to be considered did improve the average number of correct responses by one to five items per group, with most groups improving by two items. An advantage to having both strict and lenient scoring criteria is that poorer performance on certain items is more likely to be related to item difficulty or language difficulty, as the lenient criteria takes into consideration acceptable synonyms, culturally relevant slang terms, and shortened names for the item. Adding lenient scoring improves the quality of the Naming Task because it demonstrates that although participants may not use the formal name for the item, they still know what the image is representing and can name the item using terms they are familiar with. Some items (e.g., cheetah and leopard) were given two strict scores because this image was very representative for both names, and participants may not be able to accurately distinguish a difference. Some items (e.g., necklace) were removed because there were too many possible responses, making it difficult to score the item.

Based on the quality of the image, a number of items were recommended for removal. Removal criteria was determined based on the responses provided by the participants, indicating that these items were ambiguous, and thus not a good visual representation of the item in question. Furthermore, additional items were recommended for removal as they had a large number of alternate names, making it difficult to score.

There were also large language group differences, with monolingual English participants outperforming every other language group, and the bilingual French administration group performing the most poorly of all the groups. Interestingly, the monolingual French group vastly outperformed the bilinguals in the French administration. This difference might be related to the fact that the bilingual participants were selected from the Ottawa region, which is largely English-dominant. Even though all of the bilinguals had good self-reported proficiency in both languages, the environment in which they live and work may be more English-dominant than would be expected for bilinguals in Quebec City, where monolingual French participants were selected and tested.

Finally, there were a number of items where older adults outperformed the younger adults. This finding could be attributed to generational differences (Schmitter-Edgecombe, Vesneski, & Jones, 2000), or the idea that older adults may have a greater vocabulary (Hawkins et al., 1993; Sheppard et al., 2016). There may have been a number of items that older adults, but not younger ones, have been exposed to, explaining the difference between age groups (e.g., metronome). The items where there was a very large difference between older and younger adults were not necessarily recommended for removal; however, further analysis of these

items is required to determine if the generational differences are significant enough to alter the results of the test for future participants.

Future research should seek to understand why certain language groups, primarily monolingual English individuals, outperform others, and to determine how these discrepancies can be resolved to allow for the Naming Task to serve as an appropriate tool for bilingual older adults. More analysis is required to determine which images should be removed as a consequence of the inequality between language groups and age groups. Research should further focus on data collection with monolingual and bilingual patients with mild cognitive impairment conditions and Alzheimer's disease, to test the validity of the scoring criteria.

Conclusion

The present study established strict and lenient scoring criteria for an English-French picture-naming task. The Naming Task will serve as a health service for both English and French individuals to assess cognitive impairment and can be used as a suitable alternative to the BNT. The Naming Task appears to be suitable for monolingual French and English individuals. However, results are unclear when comparing bilingual to monolingual participants. Results suggest that when possible, a bilingual administration should be used when testing English-French speaking individuals, as responses will be stated in the participant's dominant language, which is affected by their language environment.

Acknowledgements

This research was supported by an Alzheimer Society of Canada Research Grant awarded to Vanessa Taler, Laura Monetta, and Shanna Kousaie (Grant #1423). The authors declare no conflict of interest. We would like to thank Linda Garcia for co-supervising this honours project. We would also like to thank Julien Blacklock, Chloe Corbeil, Dominique Fijal, Laura Thompson, Chalice Walker, Anne-Marie Lavoie, and Maude Lemieux for their assistance with data collection, as well as Jihan Nassrallah for her assistance in developing the strict and lenient scoring criteria.

References

Adesope, O. O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80*(2), 207-245. doi:10.3102/0034654310368803

Benton, A. L., & Hamsher, K. (1976). *Multlingual aphasia examination manual*. Iowa, IA: University of Iowa.

Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition, 12* (1), 3-11. doi:10.1017/S1366728908003477

Bialystok, E., Craik, F. I., & Freedman, M. (2007). Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia, 45*(2), 459-464. doi:10.1016/j.neuropsychologia.2006.10.009

Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest, 10*(3), 89-129. doi:10.1177/1529100610387084

Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(4), 859-873. doi:10.1037/0278-7393.34.4.859

Brouillette, R. M., Martin, C. K., Correa, J. B., Davis, A. B., Han, H., Johnson, W. D., ... & Keller, J. N. (2011). Memory for names test provides a useful confrontational naming task for aging and continuum of dementia. *Journal of Alzheimer's Disease, 23*(4), 665-671. doi:10.3233/JAD-2011-101455

Craik, F. I., Bialystok, E., & Freedman, M. (2010). Delaying the onset of Alzheimer disease: Bilingualism as a form of cognitive reserve. *Neurology, 75*(19), 1726-1729. doi:10.1212/WNL.0b013e3181fc2a1c

Gollan, T. H., & Acenas, L. A. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(1), 246-269. doi:10.1037/0278-7393.30.1.246

Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society, 13*(2), 197-208. doi:10.1017/S1355617707070038

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition, 33*(7), 1220-1234. doi:10.3758/BF03193224

Grant, D. A., & Berg, E. A. (1948). A behavioural analysis of degree of reinforcement and ease of shifting to new re-

sponses in a Weigl-type card sorting problem. *Journal of Experimental Psychology, 38*(4), 404-411. doi:10.1037/h0059831

Hawkins, K. A., Sledge, W. H., Orleans, J. F., Quinlan, D. M., Rakfeldt, J., & Huffman, R. E. (1993). Normative implications of the relationship between reading vocabulary and Boston Naming Test performance. *Archives of Clinical Neuropsychology, 8*(6), 525-537. doi:10.1093/arclin/8.6.525

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica, 127*(2), 277-288. doi:10.1016/j.actpsy.2007.06.003

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia, PA: Lea & Febiger.

Kohnert, K. J., Hernandez, A. E., & Bates, E. (1998). Bilingual performance on the Boston Naming Test: Preliminary norms in Spanish and English. *Brain and Language, 65*(3), 422-440. doi:10.1006/brln.1998.2001

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., . . . Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695-699. doi:10.1111/j.1532-5415.2005.53221.x

Roberts, P. M., & Doucet, N. (2011). Performance of French-speaking Quebec adults on the Boston Naming Test. *Canadian Journal of Speech-Language Pathology and Audiology, 35*(3), 254-267.

Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology, 16*(4-6), 635-645. doi:10.1080/02687030244000220

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception, 33*(2), 217-236. doi:10.1068/p5117

Schmitter-Edgecombe, M., Vesneski, M., & Jones, D. W. (2000). Aging and word-finding: A comparison of spontaneous and constrained naming tests. *Archives of Clinical Neuropsychology, 15*(6), 479-493. doi:10.1016/S0887-6177(99)00039-6

Sheppard, C., Kousaie, S., Monetta, L., & Taler, V. (2016). Performance on the Boston Naming Test in bilinguals. *Journal of the International Neuropsychological Society, 22*(3), 350-363. doi:10.1017/S135561771500123X

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643-662. doi:10.1037/h0054651

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale - Third Edition*. San Antonio, TX: The Psychological Corporation.

Appendix A. English and French Strict and Lenient Scoring Criteria.

Item #	English Strict Score	English Lenient Score	French Strict Score	French Lenient Score
1	crown		couronne	
2	helicopter		hélicoptère	
3	barrel		baril, tonneau	
4	tiger		tigre	
5	rolling pin		rouleau à pâte	
6	spool of thread, spool	thread, bobbin	bobine de fil	ficelle, fil
7	to remove image		to remove image	
8	violin, fiddle	viola	violon	viola
9	iron		fer à repasser	
10	alligator, crocodile		crocodile, alligator	
11	pliers		pince	
12	kangaroo		kangourou	
13	to remove image		to remove image	
14	duck		canard	
15	guitar		guitare	
16	trombone		trombone	
17	well	wishing well	puits	
18	to remove image		to remove image	
19	rhinoceros	rhino	rhinocéros	
20	basket		panier	
21	lobster		homard	
22	cummerbund		gaine	
23	pipe		pipe	
24	belt		ceinture	
25	ostrich	emu	autruche	émeu
26	ottoman, hassock	foot stool	pouf	coussin
27	chest of drawers, dresser		commode, buffet	
28	ruler		règle	
29	spinning wheel		rouet	
30	asparagus		asperge	
31	candelabra	candles	candélabre	chandelles
32	leopard, cheetah	jaguar	leopard, guépard	jaguar
33	racquet	badminton racquet, tennis racquet	raquette	
34	sheep		mouton	
35	door knob	knob	poignée de porte	
36	ear		oreille	
37	boot		botte	
38	ring		bague	
39	grasshopper	cricket	sauterelle	
40	nail file	file	lime à ongles	
41	screwdriver		tournevis	
42	glasses		lunettes	
43	record player	turn table	tourne-disque	
44	anchor		ancre	
45	to remove image		to remove image	
46	pineapple		ananas	
47	nut		écrou	
48	to remove image		to remove image	
49	bridle		bride	
50	hanger	coat hanger, clothes hangers	ceintre	
51	hammer		marteau	
52	to remove image		to remove image	
53	abacus		boulier	abaque
54	eagle	hawk, falcon	aigle	faucon
55	artichoke		artichaut	
56	lightswitch	switch	interrupteur	
57	carafe	decanter	carafe	
58	eye		oeil	
59	mushroom		champignon	
60	ironing board	ironing table	planche à repasser	table à repasser
61	wrench		clé (à écrou)	

Item #	English Strict Score	English Lenient Score	French Strict Score	French Lenient Score
62	onion		oignon	
63	centaur		centaure	
64	axe		hache	
65	nail		clou	
66	squirrel		écureuil	
67	lips, mouth		lèvres, bouche	
68	mitten	mitt	mitaine	
69	cannon		canon	
70	stroller, carriage	baby carriage, pram, baby stroller	pousette	landau, pousse-pousse, carosse, carosse de bébé
71	gorilla		gorille	
72	pomegranate		pomme grenade	grenade
73	wagon		brouette	
74	tambourine		tambourin	
75	heart		coeur	
76	zebra		zèbre	
77	screw		vis	
78	celery		celeri	
79	calipers		compas	
80	stool		tabouret	
81	seahorse	horse fish	hippocampe	cheval de mer
82	bow	ribbon	noeud papillon	boucle, ruban, noeud
83	rollerskate		patin à roulette	
84	glove		gant	
85	peacock		paon	
86	vest		veste	
87	kettle	teakettle	bouilloire	
88	Bunsen burner		bruleurs busen, bec Bunsen	
89	colander	strainer	passoire, égouttoir	
90	coat	jacket	manteau	veston
91	trumpet		trompette	
92	trowel		truelle	pelle de jardinage
93	raccoon		raton laveur	
94	to remove image		to remove image	
95	salt shaker	salt	salière	sel
96	arrow		flèche	
97	accordion		accordéon	
98	green pepper, bell pepper	pepper	poivron (vert)	piment, piment vert
99	broom		balais	
100	top	spinning top	toupie	
101	pitcher, jug		cruche, pichet	
102	chisel		ciseau	
103	metronome		métronome	
104	sled	toboggan	luge, traîneau	
105	hand		main	
106	monocle		monocle	
107	thimble		dé à coudre	
108	corn	corn on the cob	blé d'Inde, maïs	
109	Clothespin, clothes peg		pince à linge, épingle à linge	
110	chicken	hen	poule	
111	harp		harpe	
112	pumpkin		citrouille	
113	watering can	watering pail	arrosoir	
114	saw		scie	
115	dragonfly		libellule	
116	pear		poire	
117	rocking chair		chaise berçante	
118	windmill		moulin (à vent)	
119	butterfly		papillon	
120	to remove image		to remove image	