DNB Occasional Studies
Vol.9/No.4 (2011)

What is a fit banknote? The Dutch public responds

DNB
Occasional Studies

Frank van der Horst, Martijn Meeter,
Jan Theeuwes &
Marcel van der Woude

De Nederlandsche Bank
Eurosysteem

Central bank and prudential supervisor of financial institutions

---

Author: Frank van der Horst, Martijn Meeter, Jan Theeuwes & Marcel van der Woude

Email: m.e.van.der.woude@dnb.nl; f.van.der.horst@dnb.nl

Aim of the Occasional Studies is to disseminate thinking on policy and analytical issues in areas relevant to the Bank.
Views expressed are those of the individual authors and do not necessarily reflect official positions of De Nederlandsche Bank.

Subscription orders for DNB Occasional Studies and requests for specimen copies should be sent to:

De Nederlandsche Bank NV
Communications
P.O. Box 98
1000 AB Amsterdam
The Netherlands
Internet: www.dnb.nl

Occasional Studies

Vol.9/No.4 (2011)

Frank van der Horst[1], Martijn Meeter[2], Jan Theeuwes[2] & Marcel van der Woude[1]

What is a fit banknote? The Dutch public responds

1 De Nederlandsche Bank NV, Amsterdam

2 Vrije Universiteit, Amsterdam

Abstract

De Nederlandsche Bank (DNB) regularly checks euro banknotes in circulation for fitness for use. For this purpose it operates banknote handling machines designed to detect all types of damage, e.g. soiling or graffiti. If a banknote's quality is below a certain threshold, DNB removes the note from circulation and replaces it by a new or fit note. To our knowledge, it has never been established how the public reacts to the various types of banknote defects. It is therefore not known whether the sorting thresholds correspond with the public's view of imperfect banknotes. This could mean that we are currently removing and destroying notes that the public still perceives as fit, or on the other hand, that we return notes into circulation that are perceived as unfit for payments.

DNB, together with the department of Cognitive Psychology of the Vrije Universiteit (VU) in Amsterdam, conducted two experiments to determine the relationship between the various types and degrees of defects in circulated euro notes and the way a sample of the general public perceives the quality of these notes. The first experiment focused on single type of note defects and the second on interaction of note defects.

One of the main findings from the first experiment is that individuals do not consider limpness (lack of crispness) and folded corners a reason for rejecting notes, even when confronted with the most serious examples of these defects. This finding is relevant given that folded corners are among the main defects on the basis of which automated sorting machines in the Netherlands reject banknotes as unfit. It signifies that the sorting threshold for folded corners can be relaxed. Another finding is that the majority of the public would reject euro banknotes with large missing parts whereas automated sorting machines would not reject these notes. This implies that the threshold for this defect may need to be tightened. Finally, notes corresponding to the sorting threshold for defects tape, graffiti, stains and soil are accepted by 75% of the public.

While the variance in responses to tears, soil and mutilation was relatively small, the response variance was relatively high for other defects. Where some individuals found the smallest defect (e.g. a scribble on a banknote) cause for immediate rejection, other individuals find banknotes with serious defects (e.g. a lot of writing) still acceptable.

Individuals do add up defects when judging a banknote, although not all defects are added up to the same degree. This is the outcome of the second experiment, which addressed the interactions between the various defects. For any combination of soil, stains or tears on a banknote, the public adds up the individual ratings of the defects. On the other hand, it hardly makes a difference for the public when a folded corner is added to another defect, especially if this other defect is large. This additivity of public perception is essentially different from the way a sorting machine is programmed, as the latter do not reject banknotes if all defects are (just) below their individual thresholds.

The current sorting thresholds do not change if there is more than one defect on a note. The public however, is significantly less satisfied about notes with two defects than about single defects, when applying the same sorting thresholds. Notes with double defects, one of which is corresponding to the sorting threshold, are only accepted by 55 to 80% of the respondents. The identical single defects were accepted by 75 to 95%, which demonstrates the additivity of defects.

When formulating standards for automated sorting it is recommended to take public opinion into account. This study shows that, if it were up to the public, banknotes should not be rejected on the basis of limpness or folded corners. Based on the situation in the Netherlands, where 160 million notes are destroyed yearly, this could result in a reduction of up to one third of the banknote replacement need. On the other hand, we recommend that sorting procedures should be modified in order to take into account the additivity of defects. The output of our current, non-additive, automated sorting does not match public expectations for notes with multiple defects. In certain cases 30 to 45% of the public would reject specific defect combinations that we currently return into circulation. This implies an adjustment of the current sorting algorithms

Table of contents

1 Introduction

1.1 Research questions

To ensure that euro banknotes circulating in the Netherlands are in adequate condition, their quality is monitored by regular checks. To this end, euro banknotes in circulation are frequently passed through the banknote sorting machines of De Nederlandsche Bank (DNB), commercial banks or other recirculating parties. These machines inspect and sort on both authenticity and defects like soiling, tears or folded corners. If a banknote's quality is found to be below a minimum threshold, the banknote is removed from circulation and replaced by a new or a fit note. Most of these minimum thresholds are set by the European System of Central Banks (ESCB). So far, these thresholds have not factored in public perception of defects.

Replacement costs for banknotes should be kept as low as possible, while at the same time the public, cash handlers and other stakeholders should be satisfied with the quality and fitness of the banknotes in circulation[1]. However, to our knowledge, it has never been investigated to which degree the public accepts the various types of defects in banknotes. Therefore it is not clear whether the thresholds as defined by the ESCB match the way individuals judge banknotes. For example, it is possible that a machine programmed on the basis of the ESCB thresholds removes a banknote which would still be considered appropriate for payment by the general public, which leads to unnecessary replacement costs. On the other hand, it is not advisable to postpone replacing the notes that the public would reject. Not only should banknotes be fit in order to clearly recognize its security features but we also want to assure that the bankotes issued by DNB are accepted by the general public as a means of payment. A good quality of banknotes in circulation adds to the trust in the euro. To which degree do the sorting thresholds used at banks match with the general public's view? And is this view uniform or are there significant differences between the views of individual members of the public on the same banknote defect? And are all types of defects considered equally important?

1 The public's opinion of the cleanliness of euro banknotes in general has been stable over the years, with 83% in 2010 being considered "clean". Source: "Euro banknotes. A study about awareness and recognition of the euro banknotes among the Dutch", Visser & Sonke, TNS NIPO, 2011.

Banknotes in circulation may have not just one defect, but a combination of several smaller or more serious defects. For example, a soiled note might show some degree of wrinkling or limpness (lack of crispness) besides. If each of these defects is below the threshold, the banknote concerned will not be rejected by DNB's sorting machine. In other words, DNB's sorting standards do not add up defects but judge each individual defect as if it were the only one on the note. As individuals, on the other hand, might add up these defects, their opinions would be more severe than the machine's. It is not known if such an additive effect exists.

Members of the public may not notice differences in various increases of a defect, e.g. the length of a tear, where the sorting machine would notice a relevant increase. Therefore, a precondition for matching the public's opinion of a used banknote with the sorting machine's setting is that the public is able to differentiate between various defects with consistency.

Searching for an answer to these questions and in collaboration with the department of Cognitive Psychology of the Vrije Universiteit (VU) in Amsterdam, DNB conducted a study to determine the relationship between nine major euro banknote defects (soil, mutilation, tears, stains, tape, wrinkles, folded corners, limpness and graffiti) and the way the general public perceives these defects. The results may serve as a basis for enhancing the efficiency of the automated banknote sorting process.

The study consisted of two experiments. In the first experiment, Experiment 1, nine defects in various sizes or degrees are investigated to determine whether, individually, they would be a reason for the public to reject notes. In Experiment 2 combinations of (various sizes or degrees of) defects are examined, as notes rarely have just a single defect. Here the main question was whether a combination of defects would be more serious in the eyes of the public, than the individual defects separately. For both experiments the public's opinion is compared with the automated banknote sorting thresholds.

1.2 **Background to banknote circulation and sorting**

1.2.1 *Sorting volume*

The number of euro banknotes in circulation in the Netherlands is 300 to 400 million[2]. Being subject to wear, these banknotes undergo a quality check each time they return to a commercial bank or another recirculating party, according to ESCB rules[3]. These recirculating parties sort around 2 billion ($2*10^9$) banknotes

2 This is an estimation, because the circulation of banknotes in the Netherlands is not known due to migration-effects.

3 Council Regulation (EC) No 1338/2001 laying down measures necessary for the protection of the euro against counterfeiting, amended in 2009 via Council Regulation (EC) 44/2009, addressed at: credit institutions, payment service providers, cash in transit companies and other economic agents (e.g. traders and casinos) supplying banknotes to the public via ATMs.

per year, bringing fit banknotes into circulation again and depositing unfit ones at DNB. In the Netherlands, the central bank is the only organisation allowed to destroy unfit banknotes. DNB destroys and replaces around 160 million banknotes each year, which is about 8% of the sorting volume. In this study, we will use the term "banks" for all organisations that check the notes for authenticity and defects.

Each of the 300 to 400 million notes in circulation in the Netherlands is sorted and checked several times per year. The average life of a euro banknote depends on the denomination. A euro 5 note will last approximately one year, while a 50 euro note will last four to five years.

1.2.2 *Defects and automatic sorting*

During their life, banknotes degrade, gradually showing all kinds of defects. For the purpose of this experiment most of the defect types distinguished within the ESCB were used. Table 1 presents a detailed overview of these defects.
Most of these defects can be detected by sensors of sorting machines used by banks. For technical reasons, a few defects cannot be reliably detected at the commonly used sorting speed of 1,000 to 2,000 banknotes per minute. This is accepted because it is assumed that by the time a note starts showing undetectable defects, the note will also have developed detectable defects, and, hence, be replaced.

The ESCB has defined sorting thresholds for most of the detectable defects. There is no published ESCB threshold for wrinkles, graffiti and limpness. For these three defect types the internal DNB threshold will be used[4]. For reasons of simplicity, this study refers to the combined set of 6 thresholds as defined by the ESCB and to the 3 internal DNB thresholds as "sorting thresholds". Appendix 1 provides a detailed overview.

4 Concerning the standard for the defect graffiti national central banks in the eurozone have made non-published agreements. For wrinkles and limpness the DNB expert's idea of what is fit or not is used.

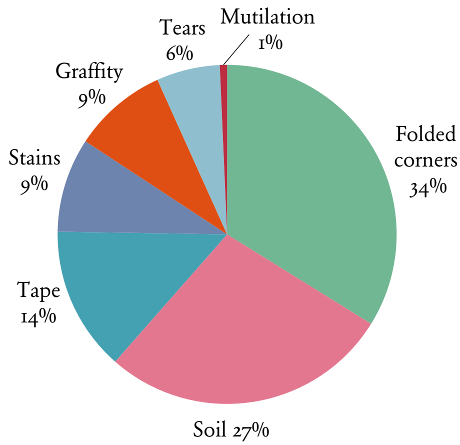
Table 1 Types of defects

Name	**Description**	**Detected by sorting machines**	**Treshold defined by (see Appendix 1 for details)**
Folded corner	Once corner of the banknote is folded, creating a so-called “dog ear”.	Yes	ESCB
Soil	Brown yellowish discoloration of the banknote, caused by handling of the note.	Yes	ESCB
Tape	Usually used on a banknote in order to repair a tear. Both dull or shiny scotch tape are used.	Yes	ESCB
Stains	Self-explanatory.	Yes	ESCB
Graffiti	Text or numbers written on the banknote.	Yes	DNB
Tears	Tears in the banknote (not repaired by tape).	Yes	ESCB
Mutilation	Pieces of the banknote have been cut or torn off.	Yes	ESCB
Wrinkles	Resulting from folding or handling the banknote.	No	DNB
Limpness	The banknote paper is less stiff as a result of frequent handling of the note.	No	DNB

1.2.3 *Frequency of defects*

Folded corners and soiling are the main grounds for rejecting notes in the Netherlands, as shown in Figure 1 below. Therefore these defects are of special interest to DNB.

Figure 1 Main grounds for destroying euro banknotes, DNB data 1st quarter 2011



2 Methods

2.1 Participants

For this study mainly visitors of the DNB Visitor Centre were invited to participate. This Centre can only be visited by groups and, commonly, such groups consist of colleagues on a social outing, members of clubs, students or families. All 42 participants in Experiment 1 were visitors of the DNB Visitor Centre. The group of 45 participants in Experiment 2 consisted of visitors of the Visitor Centre (3), DNB colleagues not working at the cash department (18) and two associations (24). For one week in October 2010 (Experiment 1) and February 2011 (Experiment 2), DNB staff selected groups of six persons willing to participate in the study.

Since these groups had diverse socioeconomic and demographic characteristics, the resulting sample was sufficiently varied. Some groups were overrepresented as compared to the general population: men (62%), those with a higher education (56%) and a higher income (22% with income > EUR 50,000 per year) and participants from the province of Noord-Holland (60%). Table 2 provides an overview of age and gender data. Moreover, it should be noted that this is not a random sample and biases may exist in the selection of the sample. In particular, visitors of the DNB Visitor Centre may be more interested in money than the average Dutch citizen.

2.2 Stimuli

Experiment 1 centred around nine single banknote defects (soil, mutilation, tears, stains, tape, wrinkles, folded corners, limpness and graffiti). The notes used had

Table 2 Demographic characteristics of the samples in Experiment 1 and Experiment 2

	Expirement 1		Experiment 2	
	Mean	SD	Mean	SD
Age	39.58	17.90	44.30	15.88
Gender	37% female; 63% male		38% female; 62% male	
Total (N)	42		45	

Table 3 Composition of banknote test sets used in the experiment

Expirement 1:	**9 Single defects:**	**5 levels**	=	**45 notes per set**
	Soil			
	Mutilation			
	Tears			
	Stains			
	Tape	**a b c d e**		
	Wrinkles			
	Folded corners			
	Limpness			
	Graffiti			
Experiment 2:	**6 combinations of defects:**	**3x3 levels**	=	**54 notes per set**
	Tears / folded corners			
	Stains / folded corners	**a+a a+b a+c**		
	Stains / tears	**b+a b+b b+c**		
	Stains / soil	**c+a c+b c+c**		
	Folded corners / soil			
	Tears / soil			

been manipulated to contain a single defect, each defect in five different levels, from level a (hardly noticeable defect), to level e (defect covering the major part of the note), with levels b – d representing approximately equal intermediate stages from a minor to a major defect (see Figure 2 for an example; and Appendix 2 for all defects). Six equal sets of banknotes were made, each set containing 45 notes (Table 3), making a total of 270 banknotes.

In Experiment 2, the respondents were offered banknotes with two defects instead of one. In order to limit the number of combinations, only four out of the nine defects from Experiment 1 were used, i.e. the ones judged by the public as the most severe. However, being only rarely encountered in practice, the defect "mutilated note" was excluded. We included the defect "folded corner", because this defect is the most common ground on which banknotes are rejected during sorting. All combinations are shown in Table 3. The largest defect levels (d and e) were not used in Experiment 2 because we expected that combinations with these levels would be rejected anyway and therefore would not be distinctive. In Experiment 2 we used six equal sets of notes, making a total of 324 banknotes.

Figure 2 Banknotes created with one of the five levels of folded corner defects (two left-hand rows) or soil defects (two right-hand rows). The defects range from minor to major in approximately equal steps. See Appendix 2 for further examples

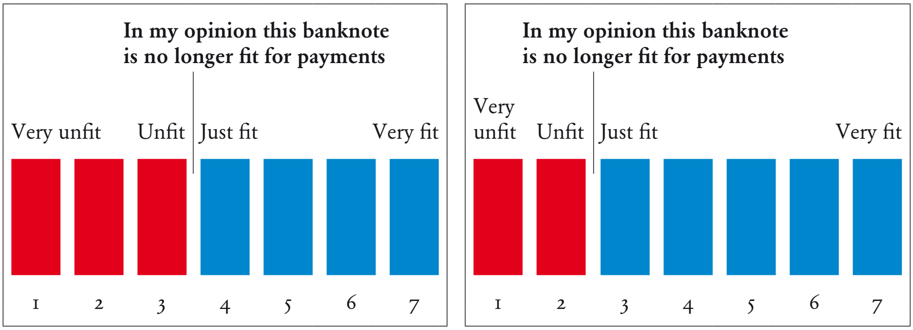
All stimuli were EUR 20 banknotes. This denomination is commonly used by the public and, at the same time, valuable enough for the public to be interested in any defects.

Not all defects were artificially created; for the defects limpness and soil, DNB operators selected banknotes from circulation having the required single defect (Experiment 1) or combination thereof (Experiment 2).

2.3 Procedure

Each test started with reading out the instructions to a group of six participants. Subsequently, each participant was given one set of banknotes and invited to place each note from the set on one of seven fields on an A1-size paper sheet, representing a 1-7 interval scale, scale point "1" denoting "very unfit for payments" and "7": "very fit for payments". The point where the scale changed from "just fit" into "unfit for payments" was indicated by both a line and a colour difference (see Figure 3). For 22 participants in Experiment 1 the scale included two fields for unfit banknotes. For the other 20 participants in Experiment 1 and for all participants in Experiment 2, the scale included three fields denoting unfit for payments (as in Figure 3). These differences in scales were introduced to test whether the scale would influence which banknotes would be rejected and which not. Appendix 3 contains pictures of the test in progress, as well as the participants' questionnaire.

Figure 3 Design of the A1-size paper sheets on which the notes had to be sorted



3 Results

3.1 **Consistency**

3.1.1 *Participants' consistency*

To assess whether participants were consistent in their ratings, in Experiment 1 we determined for each participant and each defect whether a particular rating for a given level of defect (say a rating of "4" for mutilation level b) would be reversed by a lower rating (say rating "3") if the note had a stronger defect (i.e. mutilation level c). If such a reversion of order occured relatively often it would mean that individuals are not able to rate banknotes consistently. Stated differently, if this occurred, it would follow that the banknotes' condition had no consistent effect on the subjective ratings.

The results showed that consistency was in fact quite high: 87% of the ratings were either higher than or equal to the ratings of the lower level defect, 58% of which were rated the same as the level below it. Each participant gave more consistent ratings (i.e. a higher score for a note with a less serious defect) than inconsistent ones (i.e. a lower score for a note with a less serious defect).[5] The type of question (2 or 3 reject fields) did not affect rating consistency.

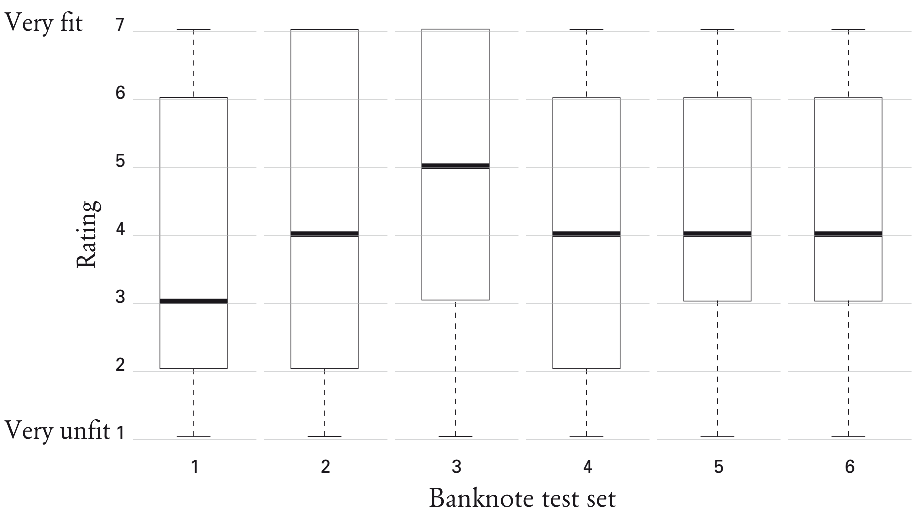
On the basis of these results it can be concluded that the participants were consistent in rating the banknotes. The high number of equally rated banknotes is not indicative of inconsistency, but may be attributed to the limited resolution provided by a 1-7 scale, or reflect that for some defects two adjoining levels of defect were too similar to merit a different rating (e.g. a 1cm long tear and a 2cm long tear could both be experienced as 'a small tear').

3.1.2 *Set consistency*

As noted, different sets of equivalent stimuli were composed manually. Even though the goal was to make these sets equivalent, minor differences between the sets remained inevitable. To determine whether this influenced the results, we checked whether ratings were different between the 6 sets in Experiment 1. An analysis of

5 A chi^2-test showed that the difference between the number of consistent and inconsistent ratings was significant for most participants in Expt 1 (22) but not all, due to a large number of equally rated banknotes

Figure 4 Score distribution for different sets in experiment 1



Distribution of ratings assigned to banknotes in each of the six sets used. The median rating is denoted by the thick black line in the box. The box gives the first-to-third quartile interval (half of all data), while the lines denote the maximum and minimum ratings given (equal to the scale limits, i.e. 1 and 7, for all sets). Set 1 was given lower ratings than the other sets.

variance with 9 (defects) x 2 (Scale type) x 5 (Level) x 6 (set) showed a main effect of set ($F = 13.459$; $p < 0.01$), suggesting that the sets were slightly different (see Figure 4). However, as no interaction with any of the other variables exists, this is an overall effect that is not related to a specific defect.

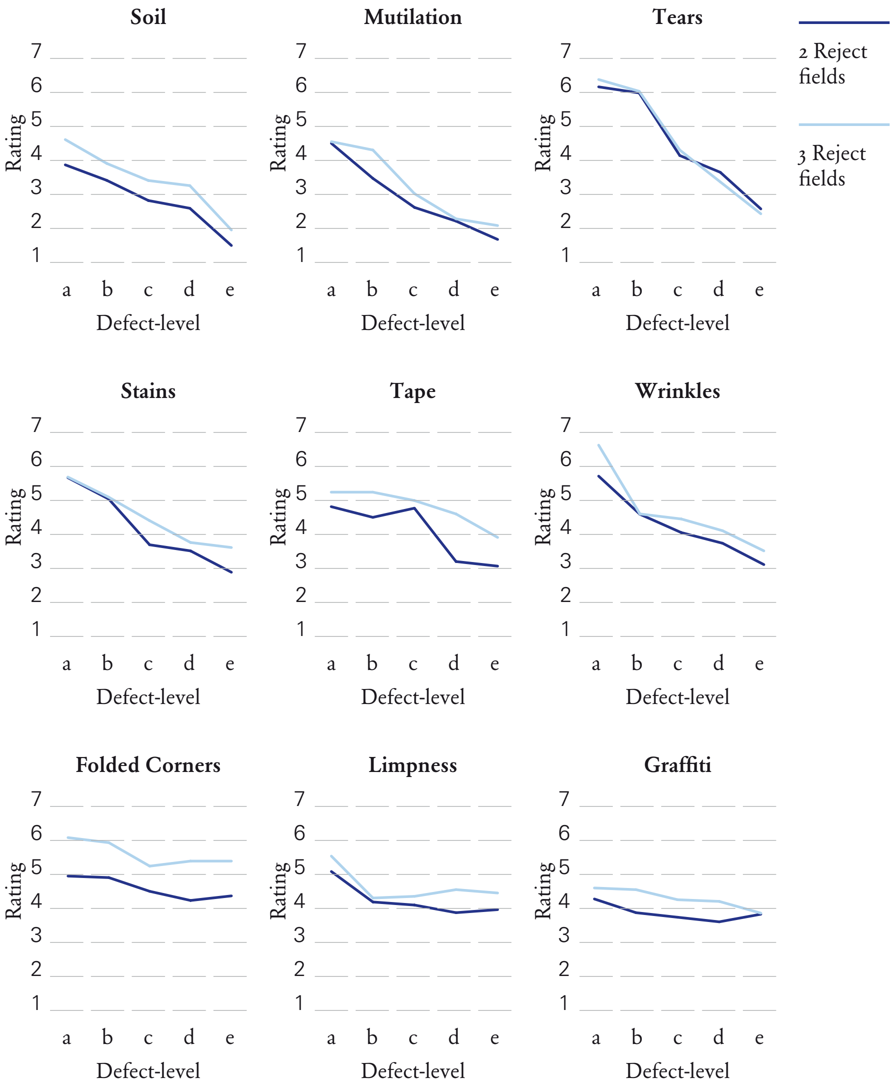
3.2 **Ratings given to single defects, Experiment 1.**

In order to gain insight into the public's opinion of notes with a single defect, the average rating at each defect level was computed. These data will allow us to measure if the public rates defects consistently and uniformly, and if there are differences in the subjective experience between defects.

3.2.1. *Rejection consistency*

Figure 5 shows the rating for the two different scales used, at an increasing defect level. As is clear from this figure, a higher level of defect correlates negatively with the score on the rating scale. This shows that individuals were indeed able to subjectively rank the level of defect in a consistent way.

Figure 5 Average scores of banknotes by defect and level in Experiment 1



Defect level e is the largest defect. Results are given separately for participants that had two of seven fields defined as unacceptable ('2 reject fields'), and those that had three fields defined that way ('3 reject fields').

Figure 6 shows the same data in a different way. Now, each rating was coded as either acceptance or rejection of the note, dependent on whether the participant chose one of the fields labelled "fit" for the note, or one of the fields labelled "unfit" (i.e. fields 1,2 and 3 in the left-hand panel, and fields 1 and 2 in in the right-hand panel of Figure 3). Consistent with the analysis of mean ratings, the higher the levels of defect, the higher the banknote rejection rate was.

Whether there were 2 or 3 rejection fields, influenced average ratings substantially (Figure 5), but not the rejection percentages (Figure 6). Both these findings indicate that participants rate the banknotes on a scale relative to the rejection point, instead of on an absolute scale. This means that opinions as to whether a note is acceptable or not are relatively independent of the scale used for the ratings. For this reason, just one scale was used in Experiment 2, i.e. the one shown in the left-hand panel of Figure 3.

3.2.2 *Rating per defect*

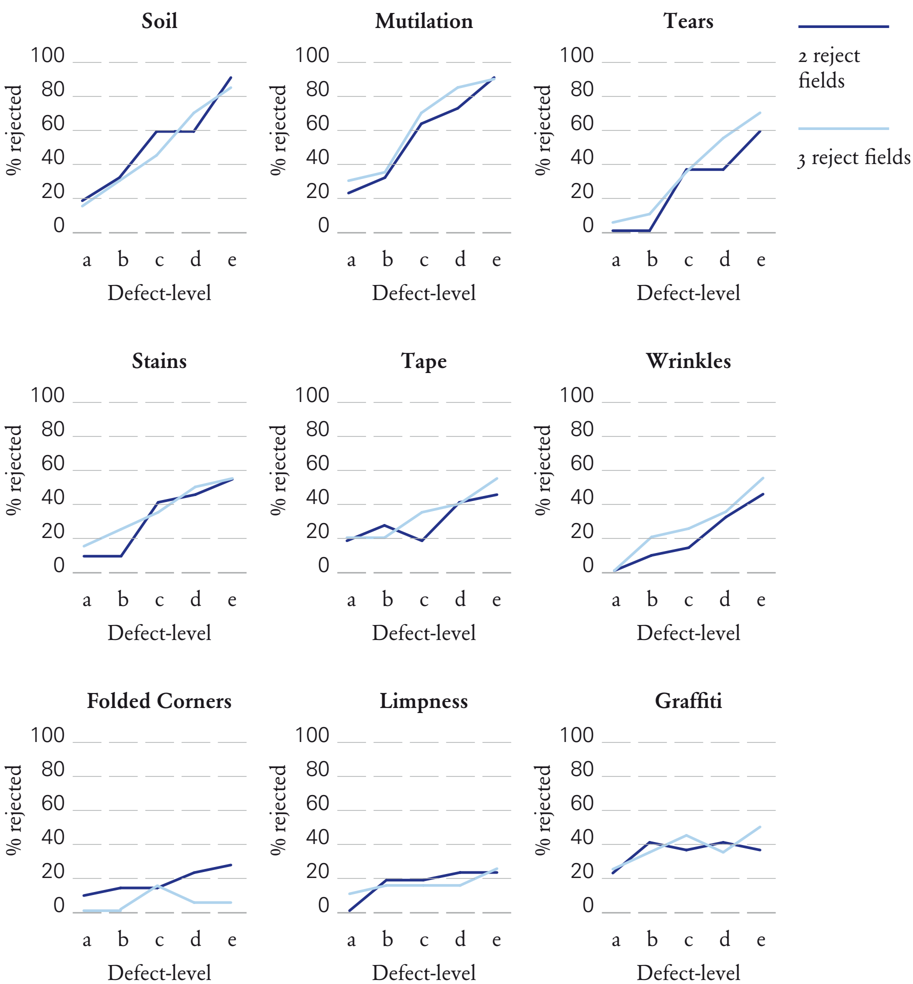
If we look at figure 6, we can discern two different reactions to defects:

- For the majority of the defects, the percentage of respondents that will reject a banknote clearly increases with the defect level. This behaviour is perhaps what would be expected, and for these defects it could be said that “Size does matter”.

- However, for notes with folded corners, limpness and graffiti, we see a different behaviour. Here the rejection percentage hardly changes with increasing defect level. So for these defects, respondents can be divided into two principal categories that either accept or reject these defects regardless of their size. All notes with graffiti were accepted by 60 to 80% of the participants. For notes with folded corners and limp notes, this acceptance is very high; at least 80% of the public is not concerned about this defect. Therefore, it seems that it would not make sense to reject and replace notes that contain one of these defects during automated sorting. Before doing so, it should be assured that folded corners do not have a negative impact on processing and distribution done by other parties, for instance on the height of the stacked notes in automated teller machines.

Furthermore, we suspect that country-specific circumstances play a role in the rating of defects: the mutilated and dog-eared test notes may at first sight appear similar, as they both miss a corner. Yet if we look at the rating given to these defects, the public differentiated their opinion of these banknotes appreciably. Firstly, this signifies that the respondents are indeed able to see subtle differences when looking at banknote defects (at least in this experiment). Secondly, the high rejection of mutilated notes could be prompted by the recollection of the old rules for mutilated Dutch guilder notes, which provided that the amount reimbursed in exchange was proportional to the remaining surface. For euro banknotes, the full value of the note is reimbursed as long as more than 50% is present. Therefore, it could very well be that specific ratings per defect would differ significantly between countries.

Figure 6 Percentages of banknotes rejected by participants



Percentages of banknotes rejected by participants (i.e. are assigned a rating labeled "unacceptable"), for 9 defects and increasing defect level, level a being the smallest and level e the largest defect.

3.2.3 Variance in opinion

The variance in our respondents' opinions of defects was relatively small for tears, soil and mutilation, but relatively large for other defects. In the latter case, the smallest defect (e.g., a bit of writing on a banknote) resulted in immediate rejection by some individuals, while for other individuals the largest defect (e.g., a lot of writing) was no reason for rejection.

3.2.4 *Comparing the public's opinion with sorting thresholds for single defects*
For a central bank it would be of interest to know if the used sorting thresholds correspond with the public's opinion. For the latter the reject percentages as shown in Figure 6 were averaged, because the use of 2 or 3 rejection fields made no difference for the rejection (see paragraph 3.2). The resulting graph (Figure 7) will now allow us to plot the sorting thresholds used for each defect during automated banknote processing, and determine the corresponding acceptance rate of the public at this threshold.

As said before, for 3 out of the 9 defects we used the internal DNB thresholds, for lack of an ESCB threshold Furthermore we employed the (arbitrary) criterion that at least 80% of the public must be satisfied, and no more than 20% dissatisfied if all notes exactly meet the sorting thresholds (green in table 4 and further. Cases where between 20 and 40% of the respondents would not be satisfied are marked yellow). Figure 7 shows both the threshold as well as the rejection rate per defect and per defect level.

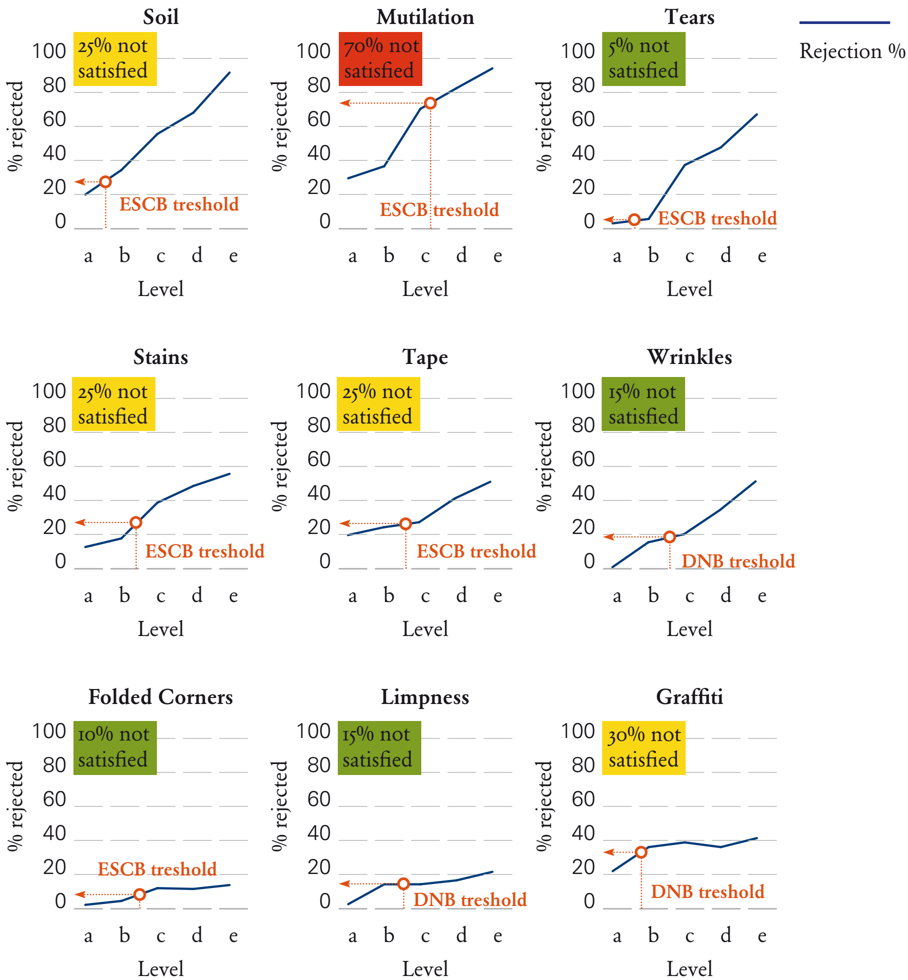
The next table outlines the results of the comparison between public perception and sorting thresholds as reflected in Figure 8. It is concluded that notes with a tear, folded corner, wrinkles or limpness at the sorting threshold level would be rejected by just a few individuals. For mutilated notes, however, this is different, because only 30% of the public is satisfied with notes that match the ESCB sorting threshold. Even though mutilated notes form only 1% of the notes rejected by DNB, it would seem that the sorting thresholds should be made more stringent.

On the other hand, for folded corners it is possible to increase the sorting threshold in line with the maximum defect level in this study, without affecting the percentage

Table 4 Percentage of public rejecting defects at the sorting threshold

Tears	5%
Folded corners	10%
Wrinkles	15%
Limpness	15%
Stains	25%
Tape	25%
Soil	25%
Graffiti	30%
Mutilation	70%

Figure 7 Applying the sorting threshold to the measured rejection percentage



The black line shows the percentage of banknotes rejected by participants (i.e. rated as unfit for payment) for increasing levels of a single defect (Experiment 1). The coloured text shows the percentage of participants that would label the banknotes “unfit for payments” at the sorting threshold. The following colours were used: green, i.e. 20% or fewer of the respondents are not satisfied; orange, i.e. between 20% and 40%; red: more than 40% is not satisfied. The red circle indicates the ESCB fitness threshold, or, if this is not defined, the fitness threshold as defined by a DNB-expert.

of the public that would still be satisfied with the quality (80%). As folded corners are among the main grounds for rejection, increasing the number of dog-eared notes returned into circulation would reduce the banknote replacement need and, hence, replacement costs. For tears we would not recommend to relax the sorting threshold given that because the public's rejection rate rapidly rises as the tear size increases.

For limpness defects, no sorting technique is available yet. We can conclude that this experiment shows there is no need to develop such a technique, because around 80% of respondents is not worried about limpness, even the level e variety.

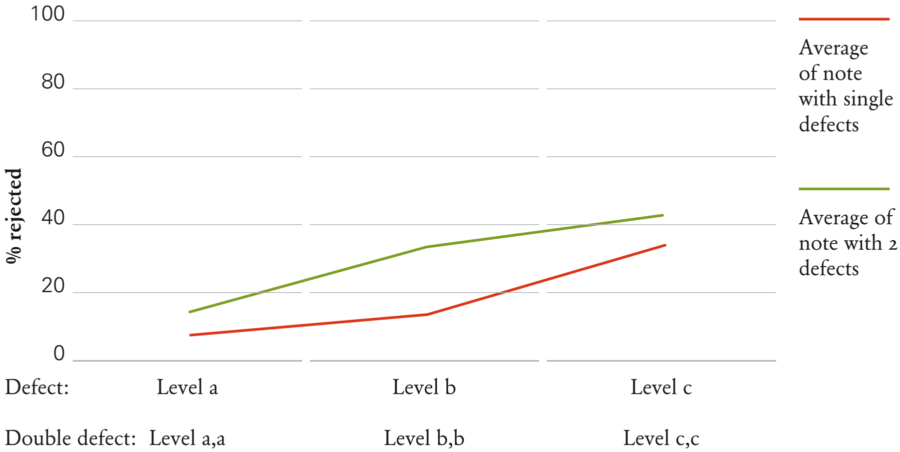
3.3 **Ratings assigned to combinations of defects, Experiment 2**

3.3.1. *Introduction*
In Experiment 2, the main question of interest is whether a double-defect banknote is judged the same as two single-defect banknotes that together have the same defects as the double-defect note. A simple hypothesis is that participants just sum up the subjective experiences of unfitness caused by the two defects. This would mean that defects affect the opinions of participants additively. On the other hand, during automated sorting, each defect is assessed separately by the sorting machine. If a particular note has two defects, and both defects are just below the fitness threshold, the note is declared fit, and will be returned into circulation. In other words, sorting machines are not programmed to "add up" defects. If there is a difference in additivity between sorting by the public and automatic sorting, this may be ground for revising the way banks should ideally operate their sorting machines.

For this experiment, the defects soil, tears, stains and folded corners were combined in six pairs. Only three of the six combinations contained notes that had exactly two defects, namely folded corners and stains, folded corners and tears, and stains and tears. These were the defects that could be artificially created on new banknotes and did not have to be selected from banknotes in circulation. For the other three combinations featuring the soil defect, it proved impossible to find banknotes in circulation that had only two defects without having, at a minor level, a third or fourth defect as well.

3.3.2. *Comparing notes with single defects and multiple defects*
Firstly, we will compare the average percentages of respondents that would reject or accept defect pairs from Experiment 2 with the average results for the single defects from Experiment 1.
Figure 8 shows the rejection percentage of notes with a pair of level a defects, a pair of level b defects and a pair of level c defects (green line). These percentages are compared with the rejection percentages at level a, b and c for single defects (red line).
From this figure, we clearly see that double-defect notes are judged more severely than single-defect notes, especially in the case of a level b defect size. It follows that, while apparently not purely mathematical, there is an additive effect.
Secondly, in order to asses this additive effect in more detail, we will focus on the results of Experiment 2 alone, in order to rule out unintentional differences in the test sets and/or participants between Experiment 1 and 2.

Figure 8 Rejection rates for single-defect notes and double-defect notes

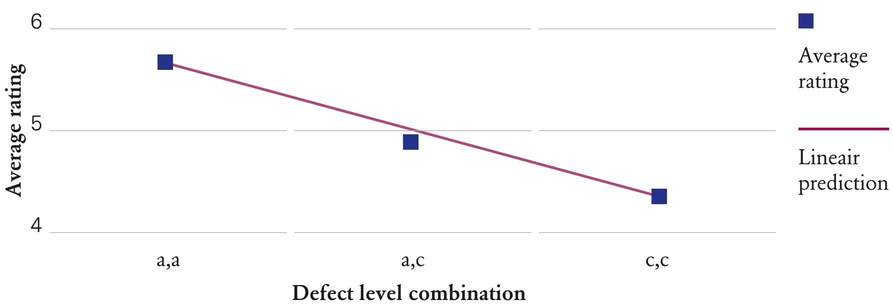


3.3.3 *Additivity: Average values*

We will investigate to what degree the public adds defects by comparing various opinions from Experiment 2. We will first concentrate on the three combinations of defects in banknotes with exactly two defects.

The clearest prediction emanating from additivity is that a note with one major defect (level c) and one minor one (level a) will be judged the same as the average of a note with two minor defects (both level a) and two major ones (both level c).

Figure 9 Predicted and observed rating for defect combination a,c



Linearly predicted rating and observed average rating given to a banknote with a combination of two defects, as a function of the level of the two defects (defect pairs on X axis; average rating on Y axis)

For this, a linear prediction of opinions is created, based on all opinions from the subjects (See Figure 9). In the same figure we have also plotted the rating of level a,c defect pairs, and that of the level a,a and level c,c defect pairs.
The blue dot at "a,c" represents the average rating for a combination of level c and a, and is slightly below, but very close to, the linear prediction. This would mean that individuals do add up several single defects in a note in a straightforward mathematical way. However, a slightly different result will appear if we look at the individual combinations of defects instead of the average values in the next paragraph

3.3.4. *Additivity: Individual Combinations*
We were especially interested to see if there was a difference between, e.g. opinions about the defect combination level c tear / level a stain versus the pair level c stain / level a tear. For each pair, the results are given in Figure 10.

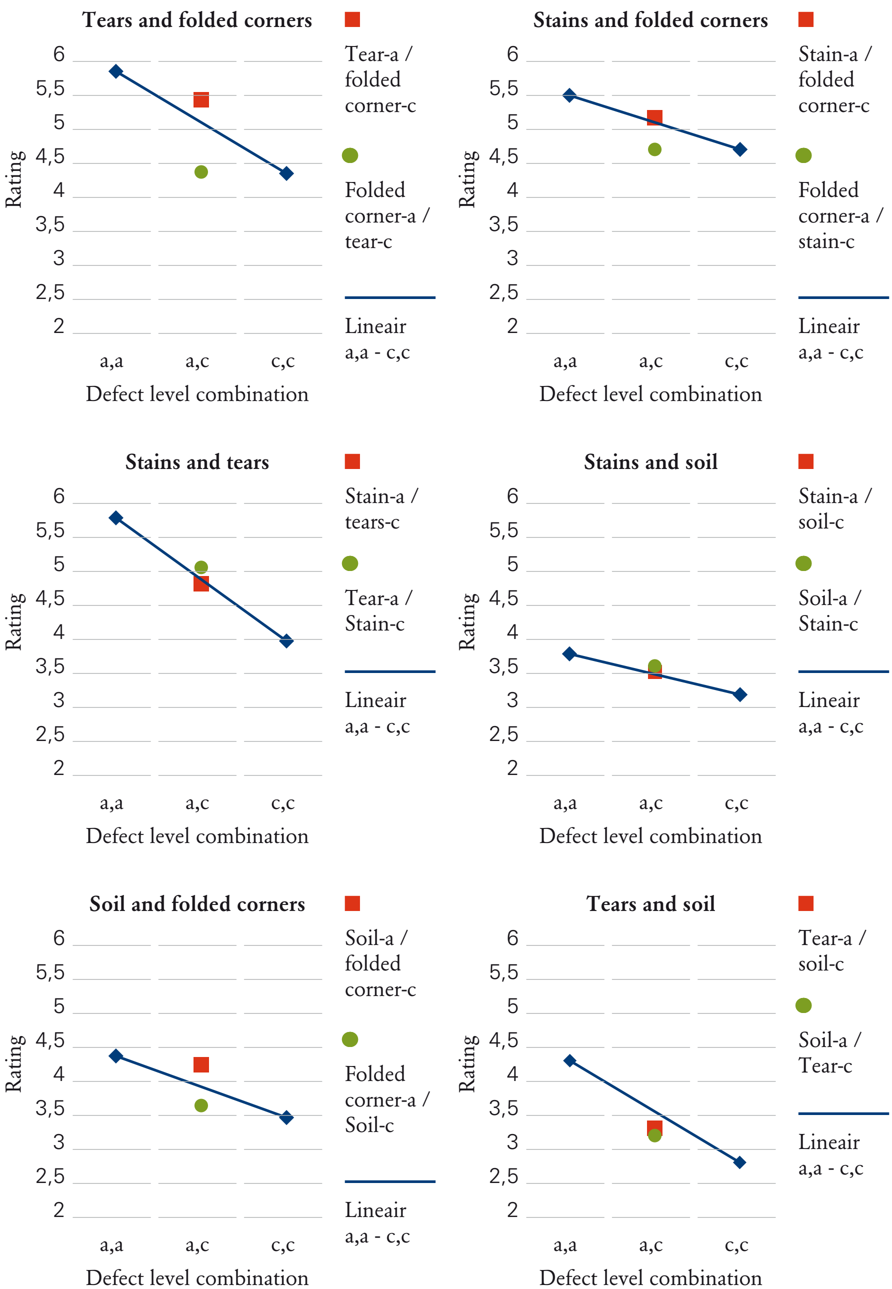
1) For all defect combinations of tears, stains and soil the opinion of the level a,c pairs matches the average of the level a,a and level c,c pairs. This means that (i) these defects behave additively and (ii) the individual defects contribute equally to opinions, because a pair with, e.g. tear level a / stain level c is judged the same as the reverse combination.

2) All combinations with folded corners are not, or only partly additive. If the other defect is a large defect (level c), increasing the size of a folded corner does not change the rating. If the other defect is small (level a), we see only a modest influence on the rating when adding a large folded corner.

We can conclude that the way participants judge combinations of defects is additive for combinations of the three defects tears, stains and soil. During automated sorting, it would be advisable to add up these defects, in order to follow the opinion of the public more closely. This entails a modification of the current automatic sorting practices.
For folded corners there is hardly an additive effect. This suggests that the current method of automated sorting, which focuses only on defects that cross the sorting threshold, is a valid approximation of how participants judge banknotes for combinations with folded corners.

If we look at the results of Experiment 1 we notice that for folded corners, limpness and graffiti the rejection or acceptance of the public did not change much with increasing defect size. So on the basis of this observation, we expect that the additive effect of limpness and graffiti is as limited as that of folded corners. For combinations of the other six defects in Experiment 1 we would expect that they will all behave additively, considering that an increasing defect size has a clear effect on ratings. This could be demonstrated in further study.

Figure 10 Defect pair ratings (level a,a and level c,c) compared with actual level a,c and level c,a defect pairs.



3.3.5 *Comparing the public's opinion with sorting thresholds for two combined defects*

We saw that for notes with single defects, the sorting thresholds were reasonbly in line with the opinion of the public, with the exception of mutilated notes. However, the sorting thresholds that we use are valid for individual defects not for combinations thereof. This means that a double-defect note will only be regarded as unfit if one of the individual defects reaches the sorting threshold. Because individuals do add up certain defects, the question is if sorting thresholds follow the opinion of the public when looking at pairs of defects.

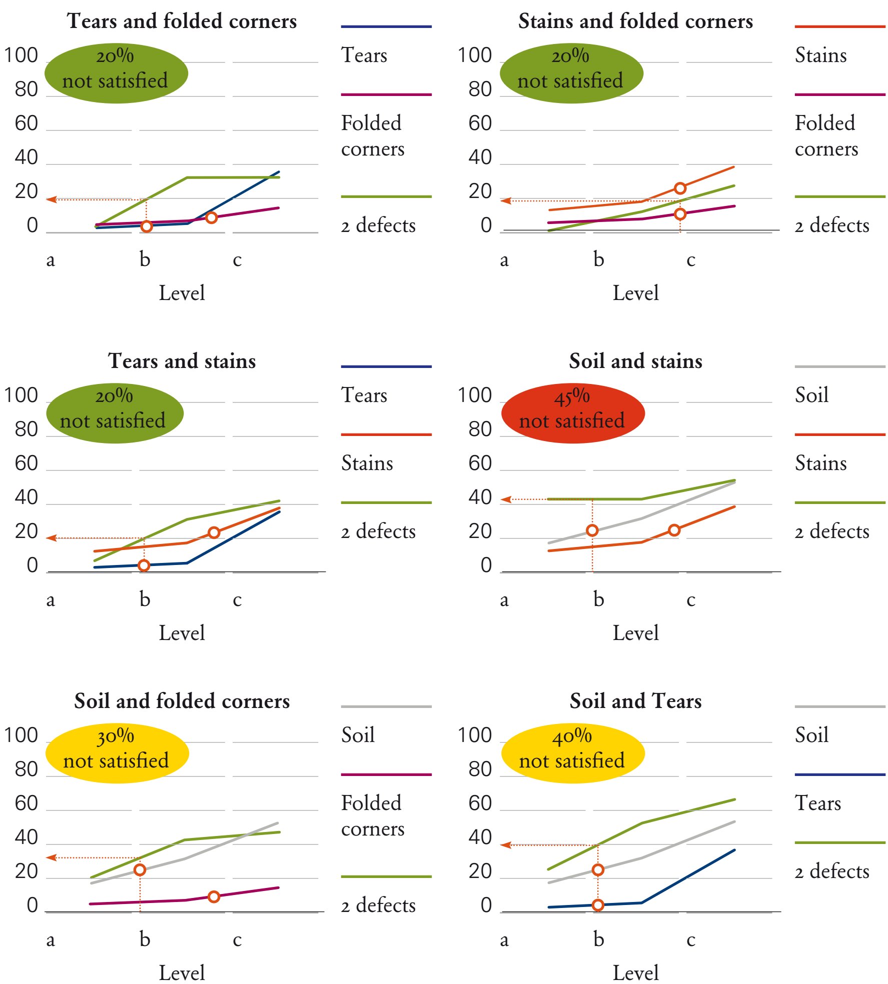
Figure 11 shows the rejection percentages for notes with 2 defects of equal size, and the rejection rate for the single defects from Experiment 1. The fitness thresholds of the individual defects have also been plotted. From this graph, we can see at what size the note with the defect pair would be rejected as a result of the sorting thresholds, and determine which percentage of the public would reject this note at this defect size. For all combinations, except the pair stains / folded corners, the rejection percentages for combinations are higher than for single defects. For the pair stains / folded corner, the rejection percentage appears to be in between the two individual defects.We could not establish if this is a valid result or an effect due to unintentional differences between experiment 1 and 2.
The results are summarised in Table 5.
The single defects stains, soil, tears and folded corners in Experiment 1 would be ground for rejection for 5 to 25% of the public, if they exactly matched the sorting thresholds. For the pairs in Experiment 2, we see that notes with combinations of the four defects would be rejected by 20 to 45 % of the respondents. This is another demonstration of the additivity of defects. Furthermore, it shows that principles for automated sorting should be changed, taking into account additivity, in order for the public's opinion to be followed more closely.

Table 5 Percentage of public rejecting defects at the sorting threshold

Tears / folded corners	20%
Tears / stains	20%
Stains / folded corners	20%
Soil / foded corners	30%
Soil / tears	40%
Soil / stains	45%

Rejection percentages for notes containing two defects of equal size, at the earliest defect size that would prompt rejection on the basis of the sorting thresholds. Summary from Figure 11.

Figure 11 Sorting thresholds and rejection percentages of combined defects



See table 5 for a summary of results.

3.4 Participant characteristics

Can one predict, on the basis of the recorded participant characteristics, how many notes participants reject? A linear regression analysis was done on the average of the ratings given by a participant to all banknotes, using as controls: age, gender, income category, subjective ratings of the importance of cleanliness and money, fear of touching dirty items, and the experiment concerned. None of these predictors had a significant effect on average ratings. We separately looked at effects of income category, education and home province on average ratings, but again no effect was found. It thus seems that demographic variables do not capture attitudes towards banknotes very well. The regression results are shown in Appendix 4.

This analysis shows there is no indication that the opinion on the fitness of banknotes differs systematically among different groups of Dutch citizens. Therefore we suggest that the sample of respondents, albeit limited, is adequate to substantiate our conclusions.

4 Conclusions

The present study shows that there is a consistent relationship between the experimentally manipulated severity of banknote defects and the way individuals subjectively judge the imperfect banknotes. This implies that people are able to rate the quality of banknotes consistently. Whether the scale we used had two or three rejection fields appeared to have not effect on the rejection rate. Participant characteristics like age, gender, or income category did not have a significant effect on average ratings.

When looking at the subjective perception of "fitness" of a banknote for payment, the single banknote defects can be divided into two categories:

1. For the majority of defects, the proportion of the public that rejects a banknote at a certain defect level, will increase with increasing defect size.
2. For notes with folded corners, limpness and writing, the rate of rejection by the public hardly changes with increasing defect size.

Furthermore, for notes with folded corners and limp notes, this rejection rate is very low, implying that the public accepts this defect at all levels. Therefore, these defects could be ignored during automated banknote sorting. In the Netherlands, this conclusion allows for a reduction of banknote replacement costs, given that the folded corners are a major rejection criterion during automated sorting.

The opposite is true for mutilated notes, as they are rejected by 70% of the public at the level corresponding with the sorting threshold. As discussed, this rejection rate for mutilated notes might be explained by the public's recollection of the old reimbursement rule for the Dutch guilder. For a Central Bank we see two possibilities for addressing this issue, either the sorting threshold for mutilated notes could be tightened, or the reimbursement rules for the Euro could be brought to the attention of the public. Given the low amount of mutilated notes in circulation, the first option is probably more cost-effective.
The variance in the opinions of our respondents was relatively small for tears, soil and mutilation, while for other defects the variance was relatively large.

Experiment 2 shows that individuals do add up the effect of two defects on a note. In this case, too, the public came up with two different reactions:

1. Soil, stains and tears are added up by the public.
2. For folded corners, a moderate additive effect for small-size defects is found, but no additive effect for large-size defects.

As a result, the public's acceptance of double-defect notes is significantly lower than for single-defect notes. As the current sorting thresholds do not change if a note has more than one defect, we have seen that the public is less satisfied when the same sorting thresholds are applied to notes with two defects. As most notes in circulation will have more than one defect, we recommend that sorting procedures should be modified in order to take account of additivity of defects. Two examples from this study can demonstrate how the current sorting approach differs from the public opinion on defects:

1. Notes with a medium or large folded corner (from level b and upwards) are deemed acceptable by 80 to 90% of the public, yet all of these notes would be rejected and destroyed by sorting machines. Potentially this could reduce the yearly replacement need in the Netherlands by up to one third.
2. On the other hand, a note with light soiling and a small stain is rejected by 45% of the public. Yet all of these notes would be brought back into circulation by sorting machines.

This study shows that there are opportunities to improve the cleanliness of notes in the eye of the public, and to sort more cost-effectively.

5 Recommendations

On the basis of this public perception study, the following recommendations for the banknote sorting process can be formulated:

- When setting thresholds for automated sorting the opinion of the public should be taken into account
- Banknotes should not be rejected on the basis of limpness or folded corners because the public seldom sees these defects as a reason for rejection
- It seems advisable to add up certain categories of defects during automated sorting. This implies an adjustment of the current sorting algorithm
- Implementing these recommendations could save significant social costs and enhance the perceived quality of banknotes in circulation.

Suggestions for future work

- Explore if and how public perception differs with the nationality of the participants.
- Examine the influence of testing with other denominations.
- Testing all possible combinations in the same test set could expand the insight on the public perception of defects.

Acknowledgements

We would especially like to thank Richard Worms, Yolanda Klaasse Bos, M'hamed Ouarous and Hans Broeders for producing the many banknotes with defects and for patiently adjusting the sets to our comments. We are also grateful to Fred Collens for reviewing the English text, to Liselotte Nyst for her help during the experiments, and last but not least to all our colleagues for their valuable input and remarks.

Appendix 1: Minimum standards for automated fitness checking of euro banknotes

Table 1 List of sorting criteria for automated fitness sorting

Defect	Definition
1. Soil	General distribution of dirt across the entire euro banknote
2. Stain	Localised concentration of dirt
3. Graffiti	Image or lettering applied in whatever manner to a euro banknote
4. De-inked note	Note featuring partial or complete absence of ink, e.g. a washed euro banknote
5. Tear	Self-explanatory
6. Hole	Self-explanatory
7. Mutilation	Damage to a banknotes that has resulted in a missing part or missing parts along at least one edge (in contrast to holes)
8. Repair	Parts of one or more banknotes joined by tape or glue
9. Crumples	Multiple random folds
10. Limpness	Structural deterioration resulting in a marked lack of stiffness
11. Fold	Self-explanatory
12. Folded corner	Self-explanatory

further information on sorting criteria

1. **Soil**

Soil increases the optical density of euro banknotes. The following table specifies the maximum density increase of limit samples that euro banknotes may exhibit to be classified as fit:

Table 2 Optical density levels

Denomination	Maximum density increase of limit sample compared to new euro banknote	Filter
€ 5	0.06	Magenta
€ 10	0.06	Magenta
€ 20	0.08	Magenta
€ 50	0.07	Magenta
€ 100	0.07	Magenta
€ 200	0.04	Magenta
€ 500	0.04	Magenta

Euro banknotes not meeting these criteria are unfit. NCBs keep reference euro banknotes showing a soil level derived from these criteria. The densitometric measurements of the reference euro banknotes are based on the following criteria:

- Standard for density measurements: ISO 5 parts 3 and 4
- Standard for the filters: DIN 16536
- Absolute measurements: standard calibration (white tile)
- Polarisation filter: on
- Aperture: 3 mm
- Illumination: D65/2
- Background: white tile standard calibration

The density increase of a reference banknote is the highest value between the averages of at least four measurement points measured on the front and on the back of the banknote in the unprinted area and without any watermark modulation.

2. **Stain**

Euro banknotes with a localised concentration of dirt covering at least 9mm by 9mm in the non printed area or at least 15mm by 15mm in the printed area are unfit.

3. **Graffiti**

There is no mandatory requirement to detect graffiti.

4. **De inked note**

Euro banknotes can be de-inked, e.g. if washed or subjected to aggressive chemical agents. These kinds of euro banknotes might be detected by image detectors or UV detectors.

5. **Tear**

Euro banknotes with tears which are open and not partly or fully covered by the machine's transport belt(s) are unfit if the tear exceeds the width or any of the lengths (depending on the tear being horizontal, vertical or diagonal) indicated below.

Table 3 Tear

Direction	Width	Lenght
Vertical	4 mm	8 mm
Horizontal	4 mm	15 mm
Diagonal	4 mm	18 mm[6]

6. **Hole**

Euro banknotes with holes which are not partly or fully covered by the machine's transport belt(s) are unfit if the hole size exceeds 10 mm^2.

7. **Mutilation**

Euro banknotes with lengths reduced by 6mm or more or widths reduced by 5mm or more are unfit. All measurements relate to differences relative to the nominal lengths and widths of the euro banknotes.

8. **Repair**

A repaired euro banknote is created by joining parts of euro banknote(s), e.g. by tape or glue. A euro banknote with tape covering an area larger than 10mm by 40mm and thicker than 50µm is unfit.

9. **Crumples**

Crumpled euro banknotes can normally be identified if their nominal level of reflectance or stiffness is reduced. No mandatory requirement applies.

10. Limpness

Insofar as possible, euro banknotes with very little stiffness are sorted as unfit. As limpness normally correlates with soiling, limp euro banknotes are generally also detected via soil sensors. There is no mandatory requirement.

11. Fold

Because of their reduced length or width, folded euro banknotes can be detected by euro banknote dimension checkers. In addition, they can be detected by thickness sensors. However, due to technical limitations, only folds fulfilling the criteria laid down for mutilations, i.e. folds leading to a length reduction in excess of 6mm or a width reduction in excess of 5mm, can be identified and are unfit.

12. Folded corner

A euro banknote with a folded corner covering an area of more than 130mm² and a minimum length of the smaller edge in excess of 10mm is unfit.

Sorting thresholds and levels of defects in the experiments

Levels a, b, c, d and e were used in experiment 1,
level a, b and c were used in experiment 2.
green: level of defect is below the threshold value;
red: level of defect exceeds threshold value

Defect	Treshold	a	b	c	d	e
Soil *	0.08	0.06	0.1	0.19	0.17	0.7
Stain	9 x 9 mm²	3 x 3 mm	8 x 8 mm	15 x 15 mm	25 x 25 mm	45 x 45 mm
Graffiti	135 mm² **	7 x 5 mm	30 x 5 mm	37 x 5 mm	55 x 5 mm	40 x 25 mm
Folded corners	130 mm²	18 mm²	105 mm²	288 mm²	392 mm²	741 mm²
Tear	4 x 8 mm	2 x 7 mm	2 x 12 mm	2 x 18 mm	2 x 36 mm	2 x 60 mm
Tape	40 x 10 mm	10 x 10 mm	30 x 10 mm	45 x 10 mm	55 x 37 mm	73 x 53 mm
Mutilation	6x71 mm***	40 mm ²	127 mm²	378 mm²	752 mm²	1666 mm²
Crumples	None ****	Good	sufficient	poor	very poor	very poor
Limpness	None ****	Good	sufficient	poor	very poor	very poor

The thresholds observed are those applying for commercial banks by ECB decision of 16 september 2010

* There is an apparent inconsistency in the values for soil for level c and d. However, when checked visually, the soil levels for these notes are all increasing, and the sorting threshold is in between a and b. In our opinion this inconsistency must be attributed to weak accuracy and capability of the density measurements for soil level determination.
** An internal ESCB threshold (referred to as ccp)
*** The threshold applies to straight sections of 5mm height and 6mm length, here converted to surface area
**** An internal DNB threshold is used

Appendix 2: Example set of banknotes used in Experiment 1

Examples of soil and dog-ear defects are given in the main text. Notes with limpness defects were not photographed. Examples of banknotes with any of the other defects are given below.

Mutilation

Stains

Teared

(A white piece of paper was inserted in the tear for visibility).

Tape

Wrinkles

Graffiti

Appendix 3: Test setting

Pictures of the test setting:

Instruction form for test of banknote appreciation by the public.

<Introduction by test instructors>

“DNB issues banknotes which initially look attractive and new but, once in circulation, are subject to wear and tear and gradually become less fit for payment. DNB operates a machine that checks banknotes for fitness. Fit notes are brought into circulation again and unfit notes are shredded. As we wish to optimise the rejection criteria of our banknote sorting machine, we are interested to find out when the general public considers circulated banknotes still fit for payment and when no longer so.

Here I’ve got a stack of notes, which as you can see are not all in the same mint condition as new notes. They are all genuine, though. We kindly ask you to place each of the notes on one of these seven fields, depending on the degree to which you still think them fit for payment.

1. If you find a banknote perfectly fit, please place it on the field on the far right-hand side.
2. If you find a banknote unfit, please place it on one of the red fields on the left-hand side.
3. If you are in doubt about a banknote’s fitness, please place it on one of the middle fields.

Please be careful not to fold or crumple the banknote, so that we know for certain that you assessed the banknote in the condition in which you received it from us. Please note that your first spontaneous response is the probably the most genuine, and therefore the exactly the one we are looking for. And remember: there are no right or wrong answers.

Do you have any questions?

After completing this test, kindly fill out the form that we will hand you in a moment."

Questionnaire pertaining to test of banknote appreciation by the public

In what year were you born?	....
Are you a man or a woman?	m/w
What is your highest educational level?	1 = elementary school
	2 = lower secondary professional education
	3 = lower general secondary education
	4 = higher general secondary education
	5 = pre-university education
	6 = higher vocational education/ university
In what province of the Netherlands do you live?	
What is your approximate gross annual income?	1 = no income
	2 = < EUR 10,000
	3 = EUR 10,000 – 20,000
	4 = EUR 20,000 – 30,000
	5 = EUR 30,000 – 40,000
	6 = EUR 40,000 – 50,000
	7 = > EUR 50,000
I avoid contact with dirty objects	1 = very unimportant
	2 = unimportant
	3 = neutral
	4 = important
	5 = very important

I love things to be clean	1 = very unimportant 2 = unimportant 3 = neutral 4 = important 5 = very important
Money is important to me	1 = very unimportant 2 = unimportant 3 = neutral 4 = important 5 = very important
Are banknotes part of your daily work routine?	Yes / No

Thank you!

Appendix 4: Regression results Participant Characteristics

Please also refer to section 3.4.

ANOVA[b] (Analysis of Variance)

Model	Sum of squares	df	Mean Square	F	Sig.
1 Regression	4,294	7	0,613	0,513	0,823[a]
Residual	95,729	80	1,197		
Total	100,022	87			

a. Predictors: (Constant), expt, gender, importance clean, age, importance of money, income, fear of touching,
b. Dependent Variable: score_mean

ANOVA[b] (Analysis of Variance)

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
1. (Constant)	9,950	16,608		0,599	0,551
Age	-0,003	0,008	-0,040	-0,318	0,752
Gender	-0,13	0,258	-0,006	-0,052	0,959
Imp. Clean	-0,206	**0,196**	**-0,136**	**-1,051**	0,296
Fear touching	-0,050	0,154	-0,0142	-0,327	0,745
Imp. Money	0,46	0,187	0,030	-0,249	0,804
Income	-3,022E-6	0,000	-0,072	-0,584	0,561
Nr. expt	0,301	0,255	-0,141	1,182	0,241

a. Dependent Variable: score_mean

Publications in this series as from January 2003

Vol.4/No.4 (2006) Monetary policy strategies and credibility – theory and practice
Bryan Chapple

Vol.4/No.5 (2006) China in 2006: An economist's view
Philipp Maier

Vol.4/No.6 (2006) The sustainability of the Dutch pension system
Jan Kakes and Dirk Broeders

Vol.5/No.1 (2007) Microfinanciering, deposito's en toezicht: de wereld is groot, denk klein!
Ronald Bosman en Iskander Schrijvers

Vol.5/No.2 (2007) Public feedback for better banknote design 2
Hans de Heij

Vol.6/No.1 (2008) Towards a European payments market: survey results on cross-border payment behaviour of Dutch consumers
Nicole Jonker and Anneke Kosse

Vol.6/No.2 (2008) Confidence and trust: empirical investigations for the Netherlands and the financial sector
Robert Mosch and Henriëtte Prast

Vol.6/No.3 (2008) Islamic Finance and Supervision: an exploratory analysis
Bastiaan Verhoef, Somia Azahaf and Werner Bijkerk

Vol.6/No.4 (2008) The Supervision of Banks in Europe: The Case for a Tailor-made Set-up
Aerdt Houben, Iskander Schrijvers and Tim Willems

Vol.6/No.5 (2008) Dutch Natural Gas Revenues and Fiscal Policy: Theory versus Practice
Peter Wierts and Guido Schotten

Vol.7/No.1 (2009) How does cross-border collateral affect a country's central bank and prudential supervisor?
Jeannette Capel

Vol.7/No.2 (2009) Banknote design for the visually impaired
Hans de Heij